VOL. XVII JUNE, 1933 No. 3

# *The Journal of* APPLIED PSYCHOLOGY

JAMES P. PORTER, EDITOR
Ohio University
WITH A BOARD OF CO-OPERATING EDITORS

## TABLE OF CONTENTS

ISABEL R. BERMAN and CHARLES BIRD. Sex Differences in Speed of Reading ..... 221

S. L. PRESSEY and L. C. PRESSEY. A Comparative Study of the Emotional Attitudes and Interests of Indian and White Children ..... 227

H. F. ADAMS. A Scale for Measuring Psychological Phenomena 239

OSCAR LITTERER. An Experimental Study of Visual Apprehension in Reading ..... 266

JESSE BATLEY RHINEHART. An Attempt to Predict the Success of Student Nurses by the Use of a Battery of Tests ..... 277

PAUL L. BOYNTON and F. A. FORD. The Relationship between Play and Intelligence ..... 294

H. S. ROGERS and G. W. HOLCOMB. An Inventory of Engineering Motives ..... 302

HAROLD V. GASKILL. Broadcasting Versus Lecturing in Psychology: Preliminary Investigation ..... 317

DONALD A. LAIRD. The Influence of Noise on Production and Fatigue ..... 320

THOMAS R. GARTH, KUNEHEI IKEDA and DOROTHY A. GARDNER. Japanese Work Curves ..... 331

NORMA V. SCHEIDEMANN. Multiplying the Possibilities of the Multiple Choice Form of Objective Question ..... 337

News and Notes ..... 341

Book Reviews ..... 344

New Books and Pamphlets ..... 347

PUBLISHED BI-MONTHLY BY THE EDITOR
Ohio University, Athens, Ohio
Printed by
C. E. PAULEY & COMPANY, INC.
225 N. New Jersey St., Indianapolis, Indiana

Entered as second-class matter at the post office at Athens, Ohio, under the act of March 3, 1879. Additional entry as second-class matter at the post office, Indianapolis, Indiana.

Acceptance for mailing at special rate of postage provided for in section 1103, act of October 3, 1917. Authorized December 21, 1922.

COPYRIGHT 1933, BY JAMES P. PORTER

# The Journal of Applied Psychology


## CO-OPERATING EDITORS

W. F. BOOK, Indiana University, W. V. BINGHAM, Personnel Research Federation, New York; ROBERT H. GAULT, Northwestern University; WILLIAM HEALY, Baker Foundation, Juvenile Court, Boston, H. L. HOLLINGWORTH, Columbia University; J. E. W. WALLIN, Atlantic University; HERBERT S. LANGFELD, Princeton University; JAMES H. LEUBA, Bryn Mawr College; HENRY M. MCCOMAS, Princeton University; WALTER D. SCOTT, Northwestern University; C. E. SEASHORE, University of Iowa; DANIEL STARCH, Boston, Mass.; EDWARD K. STRONG, JR., Leland Stanford Jr. University; L. M. TERMAN, Leland Stanford Jr. University; E. L. THORNDIKE, Columbia University; JOSEPH JASTROW, New School of Social Research, New York; G. M. WHIPPLE, Danvers, Mass.; HELEN THOMPSON WOOLEY, Teachers College, Columbia University.


The scope of THE JOURNAL OF APPLIED PSYCHOLOGY will include the following: (a) The application of psychology to vocational activities, such as law, art, public speaking, industrial and commercial work, and problems of business appeals. (b) Studies of individual mentalities, such as types of character, special talents, genius, and individual differences, including the problems of mental diagnosis and vocational prognosis. (c) The influence of general environment conditions, such as climate, weather, humidity, temperature; also such conditions as nutrition, fatigue, etc. (d) The psychology of everyday activities, such as reading, writing, speaking, singing, playing games or playing musical instruments, sports, etc.

Communications, manuscripts, reviews, summaries, notices, books and off-prints for review should be addressed to Dr. James P. Porter, Ohio University, Athens, Ohio. Contributors are requested to follow the conventional method of citation employed by this JOURNAL; in every case the year, volume and pages should be specified.


THE JOURNAL OF APPLIED PSYCHOLOGY is published bi-monthly. The subscription price is $5.00 a year. Single numbers $1.00 to $2.00, depending upon the number of pages. Remittances and business communications should be addressed to the Editor, James P. Porter, Ohio University, Athens, Ohio.

# SEX DIFFERENCES IN SPEED OF READING

ISABEL R. BERMAN AND CHARLES BIRD
*University of Minnesota*

It is scarcely necessary to review in detail the literature on sex differences in mental abilities since the field has been adequately summarized by Goodenough[1] and by Allen.[2] These reviewers agree that women are slightly superior to men in language ability and that the superiority is present at all ages. Mead[3] has found that among normal children the girls speak their first word one month earlier than the boys; McCarthy's[4] study at the pre-school level shows that girls excel in average number of words used; Pyle's[5] extensive survey of school children indicates a superiority of the girls on language tests; and Paterson and Langlie[6] report for the Iowa English test that thirty-six per cent of the freshmen boys reach or exceed the median of the girls, although sixty-one per cent of them reach the corresponding criterion in college ability. Another study at the college level is that by Eurich[7] on the reading abilities of college students. He finds that university women are superior to

[1] Goodenough, F. L. Consistency of sex differences in mental traits at various ages. Psychol. Rev., 1927, 34: 440-462.

[2] Allen, C. N. Recent studies in sex differences. Psychol. Bull., 1927, 24: 294-304; 1930, 27: 394-407.

[3] Mead, C. D. The age of walking and talking in relation to general intelligence. Ped. Sem. 1913, 20: 460-484.

[4] McCarthy, D. The language development of the pre-school child. (Institute Child Welfare Monog. Ser. No. 4) Minneapolis, Univ. of Minn. Press, 1930, pp. XII + 174.

[5] Pyle, W. H. A Manual for Mental and Physical Examination of School Children. Univ. of Missouri Bull. 1920, 21: 12, pp. 39.

[6] Paterson, D. G., and Langlie, T. A. The Influence of Sex on Scholarship Ratings. Educ. Administration and Supervision, 1926, 458-468.

[7] Eurich, A. C. The Reading Abilities of College Students. Minneapolis, Univ. of Minn. Press, 1931, pp. 60-67.

university men in a one hundred word vocabulary test but not in paragraph comprehension. In the same investigation he gives two comparisons of the sexes in speed of reading. In one group, consisting of fifty-eight women and fourteen men, the men are superior, but in another group of fourteen women and forty-nine men, the women are superior. Eurich used the Minnesota speed of reading test and had ample data to make a significant comparison; however, the results on sex differences are ambiguous because they are based on so few cases.

Apparently the question of whether women read faster than men remains open. A first clue to the presence of a difference was furnished in the data of Bird and Beers[8] on maximum and minimum inner speech. These investigators, using two forms of the Chapman-Cook Speed of Reading Test, administered one form in which subjects were to use a maximum of lip and tongue movements in silent reading and the other with a minimum of these movements. Although they did not concern themselves with sex differences in reading, it is evident that women read faster than men. The lead set in the investigation just indicated seemed to warrant an analysis of data derived from larger groups who had taken the Chapman-Cook Tests.

The present study of sex differences is based on a total of 1095 cases, 642 women and 453 men. Groups I and II, 305 subjects, were derived from the aforementioned investigation of Bird and Beers. Group I was given Form A under conditions of maximum inner speech and Form B under conditions of minimum speech. This group includes two sub-groups, Ia and Ib. Ia consisted of seventeen women and twenty-three men from the American Institute of Banking, and Ib of sixty-one women and fifty-one men from the extension

[8] Bird, C., and Beers, F. S. Maximum and minimum inner speech in reading. J. of App. Psychol., 1933, XVII, 182-187.

division of the University of Minnesota. Group II, including 101 women and fifty-two men from the summer session of the same university, read Form A with minimum speech and Form B with maximum speech. Group III was made up of 790 sophomore psychology students in the University of Minnesota, 463 women and 327 men who took the Chapman-Cook Tests as part of the Paterson and Tinker[9] studies of typographical factors influencing the speed of reading. For the comparison of sex differences, Form A alone was used since Form B was printed from a number of different type faces. Moreover, for group III scores on the Minnesota College Ability Test were available. This test is given as part

TABLE I

*Sex Differences in Speed of Reading Scores*

| | | Women | | | Men | | | | |
|---|---|---|---|---|---|---|---|---|---|
| Group | Test | N | Mean | $\sigma$ | N | Mean | $\sigma$ | Diff. | $\frac{D}{\sigma_d}$ |
| I | A | 78 | 15.49 | 4.24 | 74 | 13.05 | 3.85 | 2.43 | 3.71 |
| | B | | 17.14 | 4.01 | | 15.57 | 3.80 | 1.57 | 2.49 |
| Ia | A | 17 | 15.76 | 3.12 | 23 | 13.35 | 3.43 | 2.42 | 2.32 |
| | B | | 18.24 | 3.02 | | 15.74 | 4.33 | 2.50 | 2.15 |
| Ib | A | 61 | 15.41 | 4.50 | 51 | 12.92 | 4.03 | 2.49 | 3.09 |
| | B | | 16.84 | 4.19 | | 15.49 | 3.53 | 1.35 | 1.85 |
| II | A | 101 | 18.55 | 4.07 | 52 | 17.88 | 5.17 | .67 | .81 |
| | B | | 15.13 | 3.31 | | 14.40 | 3.67 | .72 | 1.20 |
| III | A | 463 | 19.88 | 4.07 | 327 | 18.67 | 4.43 | 1.21 | 3.92 |

[9] Paterson, D. G., and Tinker, M. A. Studies of typographical factors influencing speed of reading. J. Appl. Psychol.

of the freshman testing program at the University and is recorded in terms of percentiles of the freshman class.

In table I, which summarizes the results from these three groups, it is evident that for each group and on both forms of the test, the women are superior to the men in speed of reading. In group I, seventy-eight women read on the average 15.49 paragraphs under maximum speed conditions; whereas the seventy-four men read 13.05 paragraphs—a difference of 2.43 paragraphs in favor of the women and a $\frac{D}{\sigma_d}$ of 3.71. With minimum speech the women read 17.14 paragraphs and the men 15.57—a difference of 1.57 paragraphs in favor of the women and a $\frac{D}{\sigma d}$ of 2.49. Group II, the summer session students, gives the least evidence of a feminine superiority in reading. With Form A, minimum speech, the women read 18.55 paragraphs, or an average of .67 more paragraphs than the men; with maximum speech, the average for the women is reduced to 15.13, giving them a lead of .72 paragraphs. In both groups I and II the superiority of the women is more significant under conditions of maximum speech. This suggests either that women are more disturbed by instructions to inhibit lip and tongue movements or that they are normally inclined to use more sub-vocal speech in silent reading. Although the results so far are strongly indicative of a sex difference, they are complicated by the conditions of the silent reading experiment and by the possibility of a special selection among the men of the summer session group, since there was a larger proportion of advanced students among them. Another and unmeasured selective factor not controlled in these groups is that of differences in intelligence.

Group III permits us to check the selective influence of tested ability in speed of reading, since for this group there were available the records on the college ability

test which approximates in its construction what have been called intelligence tests. The data derived from this group are summarized in Table II. They show that women are significantly superior to men in speed of reading but not in college ability scores. The average number of paragraphs read by women is 19.88 and the average for men is 18.67. The difference of 1.21 paragraphs certainly meets the requirements of statistical significance. Although the college ability test is essentially a vocabulary test and correlates with Chapman-Cook Form A .42 for women and .52 for men, it nevertheless reveals no significant difference between the sexes. The average percentile for the women is 54.08; for the men, 51.31. This difference of 2.77 percentiles in favor of the women, in view of the low $\frac{D}{\sigma d}$ may be regarded as a chance difference.

The average woman in group III reads 1.21 paragraphs more than the average man. Her advantage in speed of reading is approximately twenty words per minute. In a comparison of individuals rather than of

TABLE II

*Sex Differences in College Ability Test Percentiles and Speed of Reading Scores in Group III, College.*

| | SOPHOMORES | | | | | |
|---|---|---|---|---|---|---|
| | 463 Women | | 327 Men | | | |
| | Mean | $\sigma$ | Mean | $\sigma$ | Diff. | $\frac{D}{\sigma_D}$ |
| C. A. T. Form A | 54.08 | 27.66 | 51.31 | 29.14 | 2.77 | 1.34 |
| Chapman-Cook | 19.88 | 4.07 | 18.67 | 4.43 | 1.21 | 3.92 |

groups, the average difference will again be 1.21, with a standard deviation of 6.02.[10] Reference to the normal

[10] $\sigma_{x-y} = \sqrt{\sigma_x^2 + \sigma_y^2 - 2r_{xy}\,\sigma_x\,\sigma_y}$

integral table shows that fifty-eight per cent of such a distribution is above zero; that is, if all possible comparisons are made between the reading speed of each woman and each man, only fifty-eight per cent of the differences will favor the women. The average differences in speed of reading, favoring the women do not seem to be a function of intelligence alone but of a general language superiority that is manifest from infancy to adult life.

# A COMPARATIVE STUDY OF THE EMOTIONAL ATTITUDES AND INTERESTS OF INDIAN AND WHITE CHILDREN

SIDNEY L. PRESSEY
LUELLA C. PRESSEY
*Ohio State University*

## 1. Nature of the Investigation

It is admittedly difficult to make any reliable comparison between two groups of people whose cultures are so different as is the case with Indians and whites. Various efforts at the measurement of comparative intelligence have been made,[1] but the upshot of these investigations to date seems to be that differences in literacy and cultural level make the usual intelligence test quite inappropriate for use with Indians and that any comparisons based on such results are inevitably unjust and invalid. In the majority at least of studies to date environmental factors have unduly influenced the results.

It would seem, however, that even though differences in intelligence may be undemonstrable—in case they exist—with the present techniques of measurement, there is still a possibility of investigating differences in attitudes, interests, or any non-intellectual traits. There have already been some scattered efforts in this field.[2]

[1] See, for instance, Garth, T. R., "The Intelligence of Mixed Blood Indians," *Journal of Applied Psychology*, 1927, IX: 268-275; Jamieson, E. and Sandiford, P., "The Mental Capacity of Southern Ontario Indians," *Journal of Educational Psychology*, 1928, XIX: 313-328.

[2] See, for instance, Kellogg, W. N. and Eagleson, B. M., "The Growth of Social Perception in Different Racial Groups," *Journal of Educational Psychology*, 1931, XXII: 367-375; Hurlock, E. B., "The Will-Temperament of White and Negro Children," *Journal of Genetic Psychology*, 1930, XXXVIII: 91-100; Young, P. C., "Intelligence and Suggestibility in Whites and Negroes," *Journal of Comparative Psychology*, 1929, IX: 339-359; Hurlock, E. B., "Color

Since the authors had available a test of such traits, developed from X-O examination,[3] it seemed of interest to study certain Indian and white groups to see what differences might appear—always reserving judgment as to the interpretation of such differences.

For the benefit of those not familiar with the tests used the following brief description may be given. The examination consists of four separate tests, each of which is merely a list of 90 words or phrases arranged in three columns. On the first test the pupils are asked to check those things they consider wrong; on the second, those things they worry about; on the third, those things they are interested in; and on the fourth, those characteristics of people that they admire. For all tests pupils may use a single check for a slight degree of feeling or a double check for a more intense attitude. The total score on the examination consists of counting, by a special weighting technique, the number of marks made by each child.[4]

The reader unfamiliar with this type of test may feel that it is just as dependent upon literacy as the usual group test of intelligence. However, there are at least two bits of evidence indicating that these tests are not as dependent as might seem. In the first place, the scores *decrease* with age—that is, the more mature the children are the fewer words they mark. If the tests measured vocabulary, the scores would presumably *increase* with age. In the second place the great majority

Preferences of White and Negro Children," *Journal of Comparative Psychology*, 1927, VII: 389-404; Stolz, A., and Manuel, H. T., "The Art Ability of Mexican Children," *School and Society*, 1931, XXXIV: 379-380; Streep, R., "A Comparison of White and Negro Children in Rhythm and Consonance," *Journal of Applied Psychology*, 1931, XV: 53-71.

[3] Pressey, S. L. and L. C., "The Interest-Attitude Tests," *Journal of Applied Psychology*, 1933, XXVII: 1-16.

[4] For precise details, see the article referred to in the above note.

of the words are simple and well within the sixth grade child's vocabulary (such words as: fire, accident, scolding, fear, brave, wealthy, storms, church, baseball). Since the test is used from the sixth grade upward, there would seem little reason to believe that the words presented could not be read by the great majority of the children. There is also no problem of connected reading; all that is required is the recognition of single words.

The purpose of the present investigation was to trace the general emotional development of sample members of the two racial groups. Every effort has been made to keep out of the picture such other factors as might influence the results. To be sure it is probable that no group of whites is exactly like any Indian group as regards economic level, social organization, and so on, but these factors have been minimized by the way in which the white schools were selected—and in any case such differences are perhaps legitimately reflected in a test of the type used.

The data have been gathered from five states: Nebraska, Montana, California, New Mexico, and Oklahoma.[5] In each state, all the children from the sixth grade up in several Indian schools were tested. Then an approximately equal number of results from white children were obtained by testing all those in grade six and above in rural communities located on, or as close as possible to, the reservations where the Indians lived. It was the authors' intention to use results from whites

[5] The writers are deeply indebted to the principals and teachers of schools attended by Indians in Macy, Winnebago, Santee, and Genoa (Nebraska); Lodge Grass, Crow Agency and Hardin (Montana); Concho and Tahlequah (Oklahoma); Shiprock (New Mexico); and Riverside (California). For the data for white children they wish to thank the principals and teachers in Macy, Waverley, Santee, Walton, Alvo (Nebraska); Lodge Grass, Crow Agency and Hardin (Montana); Calumet, Geary and Canadian County (Oklahoma); Farmington and Kirtland (New Mexico); and of all the schools in two rural counties in California.

## TABLE I

*Showing Number of White and Indian Children Tested (969 White Boys, 990 White Girls, 915 Indian Boys, 924 Indian Girls)*

| Grades | States | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Nebraska | | Montana | | California | | New Mexico | | Oklahoma | | Total | |
| 6 White | 60 | | 56 | | 92 | | 119 | | 34 | | 361 | |
| Indian | | 89 | | 30 | | 89 | | 127 | | 19 | | 354 |
| 7 White | 46 | | 34 | | 191 | | 58 | | 52 | | 381 | |
| Indian | | 74 | | 16 | | 133 | | 45 | | 66 | | 333 |
| 8 White | 35 | | 28 | | 164 | | 31 | | 70 | | 328 | |
| Indian | | 71 | | 24 | | 166 | | 18 | | 71 | | 349 |
| 9 White | 71 | | 78 | | 149 | | | | 40 | | 338 | |
| Indian | | 70 | | 19 | | 177 | | 0* | | 51 | | 317 |
| 10 White | 45 | | 65 | | 127 | | | | 30 | | 267 | |
| Indian | | 60 | | 13 | | 135 | | 0 | | 35 | | 243 |
| 11 White | 47 | | 50 | | 61 | | | | | | 158 | |
| Indian | | 40 | | 8 | | 92 | | 0 | | 0* | | 141 |
| 12 White | 49 | | 48 | | 29 | | | | | | 126 | |
| Indian | | 29 | | 15 | | 58 | | 0 | | 0 | | 102 |
| | 357 | 433 | 359 | 125 | 813 | 850 | 208 | 190 | 226 | 242 | 1,959 | 1,839 |

* Since there were no Indians tested in this state above this grade, data from the White schools have also been omitted.

who either went to the same schools with the Indians or else went to small schools in the same neighborhood. Because the Indian is typically rural it was essential that only rural whites be used, especially since the test employed shows great differences in attitudes and interests between white children living in cities and those living in the country. The whites, then, came from school systems the total enrollment of which, from the sixth grade up, did not exceed 150 children.

As a result of this work, the numbers of cases in each state were examined. See Table I. The cases are evenly distributed as to sex for all grades and both races.

It would seem that enough cases had been tested (1839 Indians and 1959 whites), and that enough care had been exercised in the selection of the white schools to be included, that comparison might yield information of general value.

## II. Results of the Investigation

The first treatment of the data obtained was to make an age-grade table for the two groups of cases. This table is presented on the next page.

It will be noticed at once that the Indian children are from one to nearly three years older chronologically, by grade, than the whites (with an average of 2.1 years). Without considering this difference in age, comparisons are likely to prove erroneous. This matter will be returned to later.

Because boys and girls tend to make quite different scores on the test used it has been necessary throughout the study to keep test results for the sexes separate. In Table III (p. 233) are presented the distribution of total scores from the Interest-Attitude Test by grade for the boys of the two races.[7]

[7] Only the distribution for the boys is presented. That for the girls may be obtained from the writers. Medians for both distributions are, however, presented later on.

## TABLE II

*Showing the Distribution of Chronological Ages for Each Grade for Indians and Whites*[6]

| Ages | Grades | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | | 7 | | 8 | | 9 | | 10 | | 11 | | 12 | |
| | W | I | W | I | W | I | W | I | W | I | W | I | W | I |
| 9 | 15 | | | | | | | | | | | | | |
| 10 | 60 | 3 | 2 | | | | | | | | | | | |
| 11 | 129 | 24 | 38 | | 8 | | | | | | | | | |
| 12 | 85 | 71 | 163 | 7 | 62 | 4 | 4 | | 1 | | | | | |
| 13 | 43 | 85 | 102 | 50 | 118 | 27 | 54 | 3 | 4 | | | | | |
| 14 | 21 | 63 | 54 | 75 | 90 | 80 | 121 | 30 | 28 | 3 | 5 | | | |
| 15 | 6 | 37 | 20 | 72 | 38 | 72 | 98 | 51 | 98 | 18 | 15 | | 2 | |
| 16 | 2 | 32 | 2 | 59 | 12 | 83 | 40 | 89 | 61 | 47 | 64 | 14 | 15 | 1 |
| 17 | | 19 | | 42 | | 42 | 15 | 80 | 57 | 69 | 57 | 30 | 53 | 11 |
| 18 | | 13 | | 21 | | 27 | 6 | 43 | 16 | 63 | 12 | 40 | 39 | 15 |
| 19 | | 2 | | 3 | | 11 | | 10 | 2 | 23 | 5 | 28 | 11 | 43 |
| 20 | | 5 | | 4 | | 3 | | 11 | | 20 | | 29 | 6 | 32 |
| Number | 361 | 354 | 381 | 333 | 328 | 349 | 338 | 317 | 267 | 243 | 158 | 141 | 126 | 102 |
| Median Age | 11.8 | 13.9 | 12.9 | 15.5 | 13.8 | 15.9 | 14.9 | 16.8 | 16.0 | 17.8 | 16.7 | 18.7 | 17.9 | 19.6 |
| Difference | 2.1 | | 2.6 | | 2.1 | | 1.9 | | 1.8 | | 2.0 | | 2.3 | |

[6] The results are not presented separately by sex because the *differences* in maturity between Indian and White boys did not differ materially from that between Indian and White girls. In both races the boys showed a slightly greater retardation than the girls. This point, however, did not seem to be important for the problem in hand. Anyone desiring results by sex may obtain the data from the writers.

## TABLE III

*Distribution of Total Score on the Interest-Attitude Tests for Indian and White Boys in Grades 6 through 12*

| | Grades | | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | 6 | | 7 | | 8 | | 9 | | 10 | | 11 | | 12 | |
| Score | W | I | W | I | W | I | W | I | W | I | W | I | W | I |
| 330-339 | | 1 | 1 | | | 1 | | | | | | | | |
| 320-329 | | 1 | | | | | | | | | | | | |
| 310-319 | | | 1 | | | | | | | | | | | |
| 300-309 | 1 | | | 1 | | | | | | | | | | |
| 290-299 | | | | | | 3 | | | | | | | | |
| 280-289 | | 1 | | 1 | | | | | | | | | | |
| 270-279 | | 3 | | | 1 | 1 | | 2 | | | | | | |
| 260-269 | 3 | 2 | 2 | | 1 | | | | | | | | | |
| 250-259 | 6 | 4 | 4 | 3 | | | | | | | | | | |
| 240-249 | 6 | 7 | 4 | 3 | 2 | 1 | | 1 | | | | | | 1 |
| 230-239 | 3 | 5 | | 2 | 2 | | 1 | 1 | | | | | | |
| 220-229 | 5 | 2 | 2 | 4 | | 6 | 1 | 2 | | | | 1 | | |
| 210-219 | 5 | 8 | | 3 | 2 | 2 | 1 | 1 | | | | | | 1 |
| 200-209 | 4 | 5 | | 6 | 2 | 1 | 2 | 1 | 1 | 2 | | | | |
| 190-199 | 10 | 5 | 4 | 2 | 4 | 5 | 1 | 3 | 1 | | | | | 2 |
| 180-189 | 5 | 7 | 6 | 4 | | 4 | 4 | 1 | 2 | | | | | |
| 170-179 | 7 | 5 | 6 | 5 | 4 | 3 | 2 | 1 | | 3 | 1 | 2 | 1 | 1 |
| 160-169 | 6 | 9 | 14 | 4 | 3 | 6 | 3 | 5 | | 2 | | | | |
| 150-159 | 10 | 8 | 6 | 3 | 8 | 5 | 3 | 6 | 1 | 1 | | 1 | | 1 |
| 140-149 | 16 | 11 | 14 | 12 | 2 | 5 | 4 | 5 | 3 | 5 | 3 | 5 | | 1 |
| 130-139 | 7 | 10 | 12 | 15 | 6 | 6 | 2 | 4 | 4 | 3 | | 1 | | |
| 120-129 | 10 | 11 | 12 | 9 | 3 | 17 | 3 | 11 | 4 | 2 | 1 | | 1 | 2 |
| 110-119 | 8 | 19 | 14 | 9 | 20 | 6 | 5 | 12 | 6 | 1 | 2 | 1 | 1 | 3 |
| 100-109 | 18 | 14 | 22 | 7 | 10 | 5 | 3 | 12 | 5 | 7 | 4 | 2 | | |
| 90-99 | 13 | 13 | 16 | 9 | 11 | 20 | 8 | 16 | 4 | 4 | 7 | 2 | 2 | 1 |
| 80-89 | 13 | 14 | 10 | 12 | 20 | 17 | 6 | 7 | 7 | 8 | 2 | 4 | | 4 |
| 70-79 | 10 | 7 | 10 | 12 | 12 | 15 | 17 | 12 | 3 | 9 | 6 | 8 | 2 | 6 |
| 60-69 | 6 | 5 | 10 | 13 | 9 | 14 | 14 | 8 | 10 | 7 | 1 | 5 | 3 | 3 |
| 50-59 | 3 | 8 | 16 | 11 | 12 | 14 | 20 | 9 | 15 | 10 | 4 | 10 | 4 | 6 |
| 40-49 | 2 | 5 | 4 | 8 | 12 | 12 | 27 | 6 | 12 | 8 | 12 | 6 | 5 | 2 |
| 30-39 | 5 | 2 | | 2 | 3 | 8 | 19 | 12 | 13 | 9 | 7 | 7 | 7 | 7 |
| 20-29 | 6 | | 2 | 3 | 4 | 6 | 8 | 5 | 19 | 8 | 6 | 12 | 10 | 3 |
| 10-19 | | | | 2 | 8 | 4 | 3 | 7 | 12 | 7 | 16 | 2 | 12 | 3 |
| 0-9 | | | | | 3 | 1 | 2 | | 4 | 1 | 3 | | 8 | 2 |
| −1—10 | | | | | 2 | | | | 1 | 3 | 2 | 1 | 2 | |
| −11—20 | | | | | 1 | | | | | | | | 1 | 1 |
| −21—30 | | | | | | | | | | | | 1 | | |
| Totals | 188 | 192 | 192 | 167 | 166 | 189 | 160 | 150 | 127 | 100 | 77 | 66 | 59 | 51 |
| Median | 130 | 128 | 114 | 115 | 88 | 91 | 60 | 95 | 52 | 66 | 44 | 54 | 26 | 65 |

In summarizing the above results for test score and chronological age, three facts have been determined for

each grade—(a) the median score, (b) the median chronological age, and (c) the median emotional age.[8] From these items, a fourth fact has been derived, (d) the years of emotional retardation. These four facts make possible a real comparison of emotional maturity between the two groups.

TABLE IV

*Showing the Median Score, Median Chronological Age, and Median Emotional Age for Indian and White Boys and Girls; also the Degree of Emotional Retardation*

| Grades | Median | | Median Chronological Age | | Median Emotional Age | | Years of Emotional Retardation | |
|---|---|---|---|---|---|---|---|---|
| | W | I | W | I | W | I | W | I |
| 6 (Boys) | 130 | 128 | 11.8 | 13.9 | 11.5 | 12.0 | 0.3 | 1.9 |
| 6 (Girls)* | 142 | 120 | 11.8 | 13.9 | 11.5 | 12.5 | 0.8 | 1.4 |
| 7 (Boys) | 114 | 115 | 12.9 | 15.5 | 12.5 | 12.5 | 0.4 | 3.0 |
| 7 (Girls) | 135 | 113 | 12.8 | 15.4 | 12.0 | 13.0 | 0.8 | 2.4 |
| 8 (Boys) | 88 | 91 | 13.9 | 16.0 | 14.0 | 13.5 | +0.1 | 2.5 |
| 8 (Girls) | 98 | 95 | 13.8 | 15.8 | 13.5 | 13.5 | 0.3 | 2.3 |
| 9 (Boys) | 60 | 95 | 15.1 | 16.8 | 15.5 | 13.5 | +0.4 | 3.3 |
| 9 (Girls) | 64 | 84 | 14.9 | 16.8 | 14.5 | 14.0 | 0.4 | 2.8 |
| 10 (Boys) | 52 | 66 | 16.1 | 18.1 | 16.0 | 15.0 | 0.1 | 2.9 |
| 10 (Girls) | 41 | 65 | 15.9 | 17.6 | 16.0 | 14.5 | +0.1 | 3.1 |
| 11 (Boys) | 44 | 54 | 16.7 | 19.0 | 16.5 | 16.0 | 0.2 | 3.0 |
| 11 (Girls) | 36 | 55 | 16.6 | 18.1 | 16.0 | 15.0 | 0.6 | 3.4 |
| 12 (Boys) | 26 | 65 | 18.0 | 19.9 | 17.5 | 15.0 | 0.5 | 4.9 |
| 12 (Girls) | 20 | 59 | 17.7 | 19.2 | 17.5 | 14.5 | 0.2 | 4.7 |

* The median scores for the girls, the derived median emotional ages, and the median chronological ages per grade have been calculated from tables not presented in detail in the article. The median chronological age for boys was also derived from the table for that sex alone.

Three conclusions are immediately clear from the above figures: (1) that the Indians are older per grade

[8] For the derivation of these "emotional ages" from the raw scores and for a general discussion of the concept, see the reference to the Interest-Attitude Tests above.

than the whites, (2) that the emotional age of the Indians is considerably behind their own chronological age, and (3) that in the lower grades the Indians make more mature scores than the whites while in the upper grades they make less mature scores. The differences in emotional maturity, as the medians now stand, without compensating for differences in actual age, show the younger Indians to have been retarded until their advanced age makes them seem to be on a par or more mature than the whites of the same grade. However, the Indians are an average of 2.1 years older per grade. When this item is taken into consideration, one finds the Indians to be actually less mature emotionally by at least one year and sometimes as many as four. If one takes for each grade the difference in chronological age and adds to it the difference in maturity (using a + sign if the difference is in favor of the white and a — sign if it is in favor of the Indians) one comes out with a series of differences such as are given below:

TABLE V

*Showing the Actual Differences in Level of Emotional Development between Indians and Whites when Differences in Chronological Ages are Taken into Consideration*

| | Grades | | | | | | |
|---|---|---|---|---|---|---|---|
| | 6 | 7 | 8 | 9 | 10 | 11 | 12 |
| Boys | 1.6 | 2.6 | 2.6 | 1.9 | 3.0 | 2.8 | 4.4 |
| Girls | 1.1 | 1.6 | 2.0 | 2.4 | 3.2 | 2.5 | 4.5 |

For instance, in Grade 6, the white boys test 0.5 years less mature than the Indians; they are, however, 2.1 years younger chronologically. Their actual excess of maturity is then 1.6 years. The sixth grade white girls test 1.0 younger emotionally than the Indians; they are, however, 2.1 years younger chronologically. Subtracting one of these differences from the other, the actual maturity becomes 1.1 years. On the other hand, the tenth

grade boys show a chronological difference of 2.0 years but, in spite of it, test 1.0 year older; adding these together, one gets an actual difference of three years.

## III. Discussion of the Results

It should once more be pointed out that the test scores decrease with age and so can hardly be reflections of either greater vocabulary or greater schooling. As near as can be determined these scores reflect what might best be called "sophistication"—knowledge of the world, freedom from childish morality and blinding worries and superstitions. It is the general concensus of opinion on the part of those dealing with them that the Indians do not become free from such attitudes—in particular, they do not seem to pass through the period of "emancipation" during adolescence as in the case with the white children. This lack of adolescent development is shown by the generally greater and greater distance between the two racial groups as the children become older. The Indian tends to remain immature; either he is incapable of a more mature adjustment or else his environment has been so simplified that adjustment on a childish level is good enough.

Those working with or interested in Indians should not in the least feel that the situation is either a criticism of the Indian or a statement of an irremedial condition. Emotional attitudes, unlike intelligence, are extremely modificable. If Indians are, generally, childish there is no reason why they have to remain that way. Presumably contact with more mature attitudes, exposure to adult ideas, encouragement of less childish interests, and the providing of more motivation for further development may result in bringing Indians up to the level of the whites in this respect. Surely, those Indians who have come into contact with white civilization through their schooling (incidentally, all those tested belong in

this category) are more mature emotionally and socially than those still living in isolation out on a reservation. This same statement would presumably hold true of isolated versus "civilized" groups of any race.

The Indian's environment makes him fear and dread such things as fires, floods, prison, murder, fights, robbers, and so on; the white person of the same age worries over appearances, manners, morals, money, clothes, a job, his family, and the like. To say that one group is less intelligent or less well adapted than the other is absurd—and presumably one group will become like the other (if likeness is desirable) in proportion as the social and economic environments become similar. It may even be that a less mature emotional development for the Indian is better, although *a priori* it would be hard to think up reasons for this point of view.

The writers would like to advance the opinion that it is in such differences of attitude and interest as are reflected by these tests that the greatest differences between races occur, and with these differences the greatest opportunity for misunderstandings. It is likely that the white man has antagonized the Indian more by his insistence upon particular social customs, personal efficiency, and so on than by his superior intelligence—in case he has it. It is possible that those in charge of Indians may be brought to a better understanding of them by means of such investigation as this—with a resulting better adjustment between the two races.

The differences thus far discussed are concerned wholly with the total score on the examination used. It is intended later to present the results term by term so that the reader may see just which ideas of right and wrong, which worries, which interests, and which admirations show slower development in one group than in the other. Presumably the gross differences are due to variations in particular items—not in every item—and these specific items probably reflect the real differ-

ences in attitudes, the totality of which is given in the above scores. A little such analysis is already completed, with interesting results. Thus, white children worry more and more about money as they grow older; during childhood the Indians worry a little about money, but by adolescence the worry disappears almost completely —being retained apparently only by those few who are not content to let Uncle Sam do the worrying. Probably also, this attitude is due in some measure to the general lack of interest in and worry about material possessions as shown by the Indians throughout the tests. These analytic results show also that white children in increasing numbers regard "efficient" as an admirable characteristic in people; but Indians, at no age, regard this trait as being admirable—in fact, one gets the impression they consider it as something of a vice! When the results by item are complete it should be possible to get a fairly accurate picture of the real sources of differences in attitude—insofar as these sources are covered by the items of the test.

# A SCALE FOR MEASURING PSYCHOLOGICAL PHENOMENA

HENRY F. ADAMS
*University of Michigan*

A conventional scale of measurement may be developed if any function of judgments of more or less is found to be correlated with differences in amount of the items judged, provided equal units of amount of difference can be determined and expressed on a linear scale. Such judgments of more or less may be obtained in a variety of ways, but the one selected for analysis in this paper is the rank difference method.[1]

The method can be validated and its true worth appraised by a series of relatively simple experiments. If, for example, a number of weights be prepared so that one shall weigh 30 grams, another 35, a third 40, and so on, and if a sufficient number of individuals arrange these items in the order of apparent weight, then a comparison between the experimental results and the actual order (the amounts of the weights) will determine the validity of the method. Such a series of experimental investigations was undertaken. Six experiments were performed.

Experiment 1. The material consisted of five series of brown cardboard pill boxes which were loaded with sand, shot and cotton to bring them to the desired weight. The boxes were cylindrical in shape, the height being 2½ inches, the diameter 1⅝ inches. Ten boxes constituted one series. The increments of difference were 1/20, 1/30, 1/40, 1/50, and 1/60, respectively. One hundred ten subjects arranged each series in order of relative weight twice, thus giving a total of 220 records.

[1] See the paper by the writer which appeared in the Journal of Social Psychology, Vol. I, 1930, p. 122.

Experiment 2. The material was composed of 5 series of 10 lines each. Each line was drawn in black India ink on a 4″ x 6″ plain white filing card. The increments of difference were 1/50, 1/75, 1/100, 1/125, and 1/150, respectively. Seventy subjects performed the experiment twice, thus giving a total of 140 reactions to each line.

Experiment 3. The material was made up of 5 series of 10 circles each. Each circle was drawn in black India ink on a plain white 4″ x 6″ filing card. The diameters of the circles were equal to one-half the lengths of the lines in experiment 2. Seventy subjects, the same as those used in experiment 2, performed the experiment twice, consequently there is a total of 140 reactions to each circle.

Experiment 4. The material consisted of 7 series of 10 lines each. Each line was drawn in black India ink on a 4″ x 6″ plain white filing card. Throughout, the fifth line of each series was 100 mm. long. In series 1, a constant difference of 5 mm. between contiguous lines was maintained; in series 2, a difference of 4 mm.; in series 3, a difference of 3 mm.; in series 4, a difference of 2 mm.; in series 5, a difference of 1 mm.; in series 6, a difference of ½ mm.; in series 7, a difference of ¼ mm. Thirty-four subjects performed the experiment twice, making a total of 68 reactions to each line.

Experiment 5. The material was composed of 7 series of 10 circles each. Each circle was drawn in black India ink on a 4″ x 6″ plain white filing card with a diameter equal to ½ the length of the corresponding line of experiment 4. Thirty-three subjects performed the experiment twice, thus giving a total of 66 reactions to each circle.

Experiment 6. The material was made up of 7 series of 10 triangles each. Each triangle, drawn in black

India ink on a 4″ x 6″ plain white filing card, was equilateral, the side being equal to the diameter of the corresponding circle in experiment 5. Thirty-three subjects performed the experiment twice, thus giving a total of 66 reactions to each triangle.

*Self consistency* may be considered as showing how well the same person agrees with himself on the two experimental occasions. *Group consistency* indicates the degree of agreement existing between two or more different individuals who perform the experiment under conditions as identical as possible. *Accuracy* discloses the agreement between the individual performances and the perfect performance.

## Results

Our immediate concern is, of course, with the validity of the rank difference method. But this is so clearly a function of all of the results that the discussion of the validity of the technique must keep pace with the particular conclusions reached.

The eight values obtained from each experimental series will now be presented. In the first column of the table is given the number of the experiment and the series number. For example, 1-1 means experiment 1, series 1; 1-2, experiment 1, series 2; 2-1, experiment 2, series 1. In column 2 is given the average self consistency for the particular series. In column 3 appears the group consistency; in column 4, the average accuracy; in column 5, the average deviation; in column 6, the standard deviation; in column 7, the error; in column 8, the range; in column 9, the percent correct; in column 10, the amount of stimulus difference. The relative amount of stimulus difference was determined in this manner. The absolute physical amounts, for example, the areas of the circles in series 1, were added together and averaged. This gave a value of 21.03 sq. cm. The average difference in area was next obtained by sub-

## FUNDAMENTAL TABLE[1]

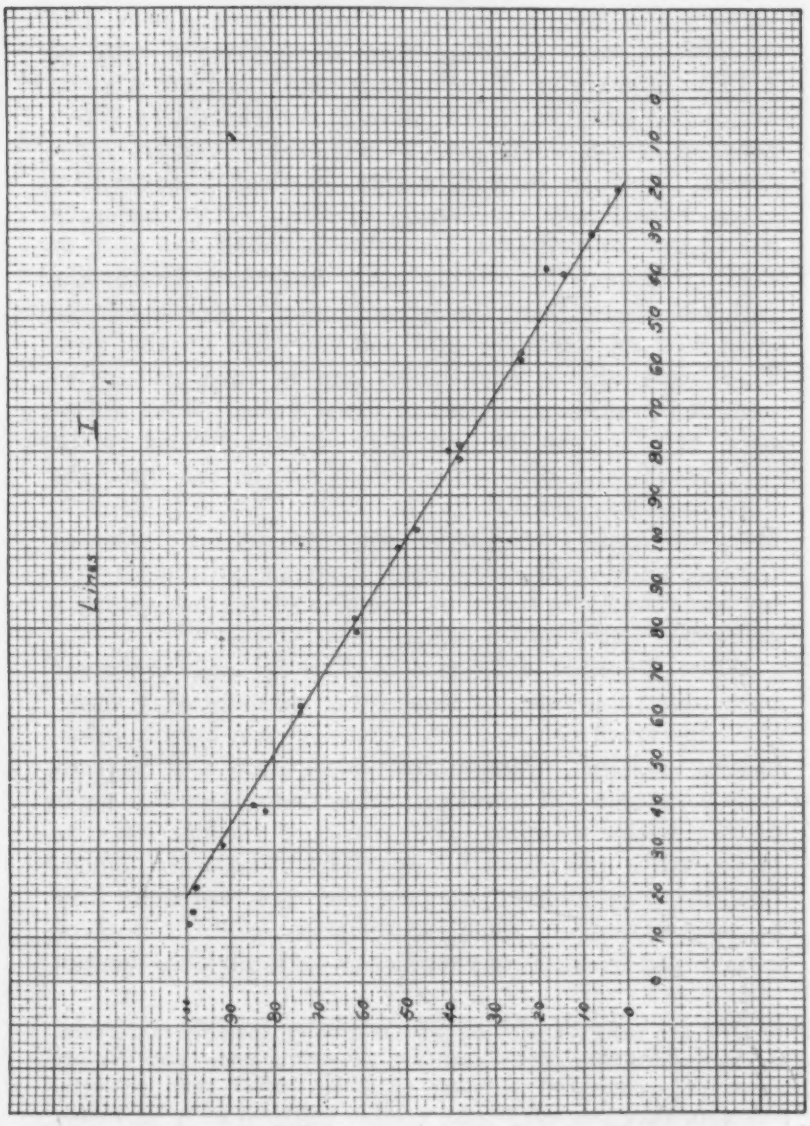
| No. | SC | GC | Acc | A D | S D | Er | Ra | %C | Unit |
|---|---|---|---|---|---|---|---|---|---|
| 4-1 | .995 | .995 | .998 | .022 | .093 | .013 | .999 | .953 | .120 |
| 5-1 | .996 | .996 | .998 | .017 | .038 | .012 | .997 | .958 | .130 |
| 4-2 | .986 | .981 | .993 | .057 | .118 | .032 | .994 | .902 | .150 |
| 6-1 | .993 | .994 | .998 | .030 | .087 | .018 | .992 | .932 | .170 |
| 5-2 | .988 | .986 | .996 | .059 | .132 | .036 | .987 | .882 | .170 |
| 4-3 | .984 | .984 | .992 | .074 | .122 | .041 | .973 | .871 | .200 |
| 6-2 | .989 | .989 | .994 | .060 | .131 | .035 | .988 | .872 | .220 |
| 5-3 | .956 | .958 | .979 | .146 | .213 | .089 | .961 | .726 | .220 |
| 3-1 | .919 | .918 | .958 | .194 | .279 | .124 | .924 | .655 | .260 |
| 6-3 | .965 | .957 | .978 | .139 | .223 | .091 | .973 | .720 | .290 |
| 4-4 | .916 | .919 | .963 | .177 | .289 | .108 | .960 | .725 | .310 |
| 5-4 | .886 | .904 | .952 | .260 | .326 | .168 | .906 | .531 | .320 |
| 3-2 | .878 | .884 | .937 | .270 | .381 | .198 | .900 | .484 | .360 |
| 2-1 | .849 | .856 | .933 | .313 | .381 | .192 | .907 | .540 | .400 |
| 6-4 | .901 | .887 | .948 | .284 | .344 | .189 | .881 | .490 | .430 |
| 3-3 | .851 | .845 | .916 | .333 | .426 | .239 | .900 | .424 | .480 |
| 3-4 | .733 | .745 | .854 | .452 | .519 | .322 | .811 | .311 | .600 |
| 4-5 | .753 | .736 | .850 | .449 | .442 | .307 | .812 | .379 | .610 |
| 2-2 | .757 | .744 | .850 | .416 | .525 | .311 | .787 | .347 | .620 |
| 5-5 | .711 | .718 | .851 | .449 | .513 | .327 | .823 | .310 | .640 |
| 1-1 | .722 | .710 | .834 | .461 | .555 | .348 | .872 | .326 | .670 |
| 3-5 | .626 | .639 | .795 | .540 | .610 | .378 | .745 | .297 | .780 |
| 2-3 | .632 | .614 | .783 | .560 | .625 | .388 | .745 | .289 | .790 |
| 6-5 | .649 | .600 | .782 | .570 | .632 | .405 | .761 | .211 | .830 |
| 1-2 | .573 | .570 | .746 | .629 | .686 | .478 | .718 | .191 | .920 |
| 2-4 | .530 | .479 | .692 | .688 | .732 | .477 | .679 | .232 | .980 |
| 1-3 | .402 | .405 | .640 | .752 | .800 | .586 | .580 | .133 | .820 |
| 4-6 | .407 | .387 | .639 | .724 | .768 | .540 | .627 | .176 | .820 |
| 2-5 | .473 | .406 | .635 | .716 | .776 | .540 | .590 | .052 | .800 |
| 5-6 | .411 | .351 | .600 | .763 | .816 | .597 | .549 | .120 | .780 |
| 1-4 | .297 | .258 | .458 | .848 | .880 | .703 | .478 | .081 | .650 |
| 6-6 | .234 | .213 | .480 | .861 | .878 | .712 | .507 | .079 | .590 |
| 1-5 | .166 | .197 | .360 | .902 | .918 | .773 | .386 | .051 | .550 |
| 4-7 | .153 | .184 | .410 | .882 | .931 | .734 | .470 | .093 | .410 |
| 5-7 | .121 | .134 | .348 | .922 | .952 | .773 | .340 | .059 | .390 |
| 6-7 | .108 | .063 | .247 | .950 | .973 | .852 | .318 | .044 | .290 |

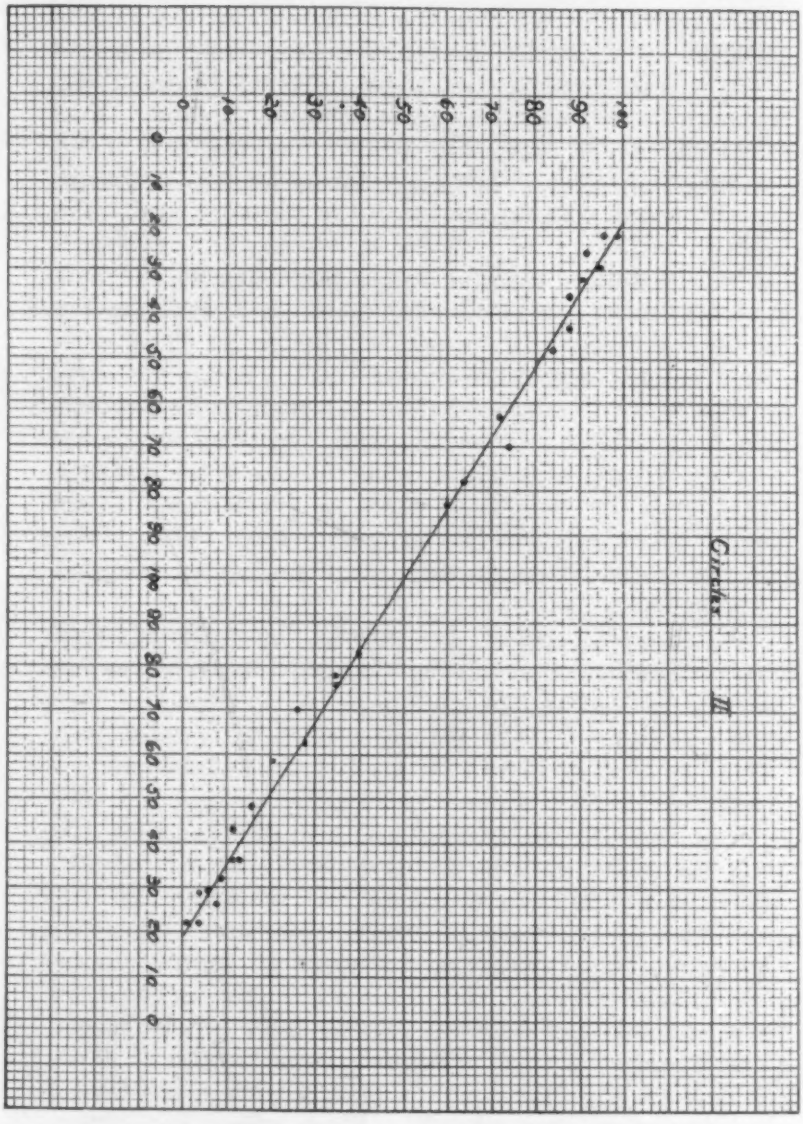
[1] The first five rows in the table are not used in the subsequent graphs, for the accuracies are all equal to 1.00 approximately. The entries between .200 and .920 are positive, for the unit is less than the difference. Those below the line, extending from .980

tracting the area of the smallest from the area of the largest, and dividing by n-1. This difference proved to be 2.00 sq. cm. Dividing 2 by 21.03 gives a ratio of 1/10.5. The series of ratios thus obtained from a given series was graphed, and that ratio determined by interpolation, which would give 75 per cent of correct judgments. This, by the unlike signs method, is equal to a group consistency correlation of 0.50, or an accuracy correlation of 0.707. This ratio is called the *unit of discrimination.* Once this ratio has been found, the ratios obtained from the experimental material can be expressed as multiples and sub-multiples of the unit. Then each of the 8 variables can in turn be plotted against the scale composed of equal units. With differences greater than the unit, the unit was divided by the difference; with differences smaller than the unit, the difference was divided by the unit. This was done to make the scale extend equal distances on both sides of the mid points. If it were not done, an arithmetic progression would be found on one side of the mid point, a geometric series on the other.

When the rank difference method is used any error must be accompanied by some other compensating error. Such material, consequently, constitutes a closed, interlocking system. Change introduced into any part of such a system will affect the interrelations of the parts of the entire system. The resulting change may be measured in any of the 8 ways described above. Therefore, it is reasonable to suspect that these variables may be interrelated. If such interrelation exists it may be determined by plotting graphically each variable against

to .290, are negative, for the unit is greater than the difference. Average deviation, range, standard deviation, error and percent correct are expressed in terms of ratios rather than percents. For example, when n is 10, the average deviation given by chance is 2.50. The obtained A. D. is divided by 2.50 in the table above. Standard deviation and error are treated in the same way.


T
Years
0
10
20
30
40
50
60
70
80
90
100
110
120
130
140
150
160
170
180
190
200
210
220
230
240
250
260
270
280
290
300
310
320
330
340
350
360
370
380
390
400
410
420
430
440
450
460
470
480
490
500

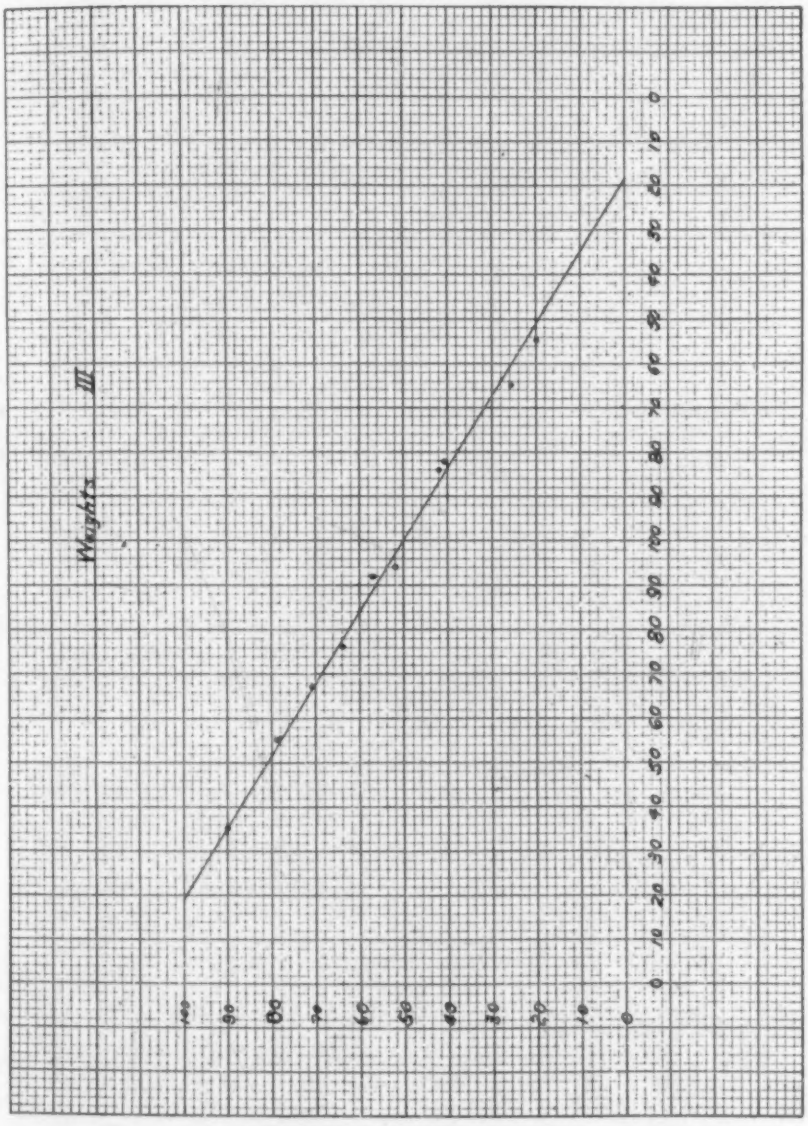

0
10
20
30
40
50
60
70
80
90
100
Circles III

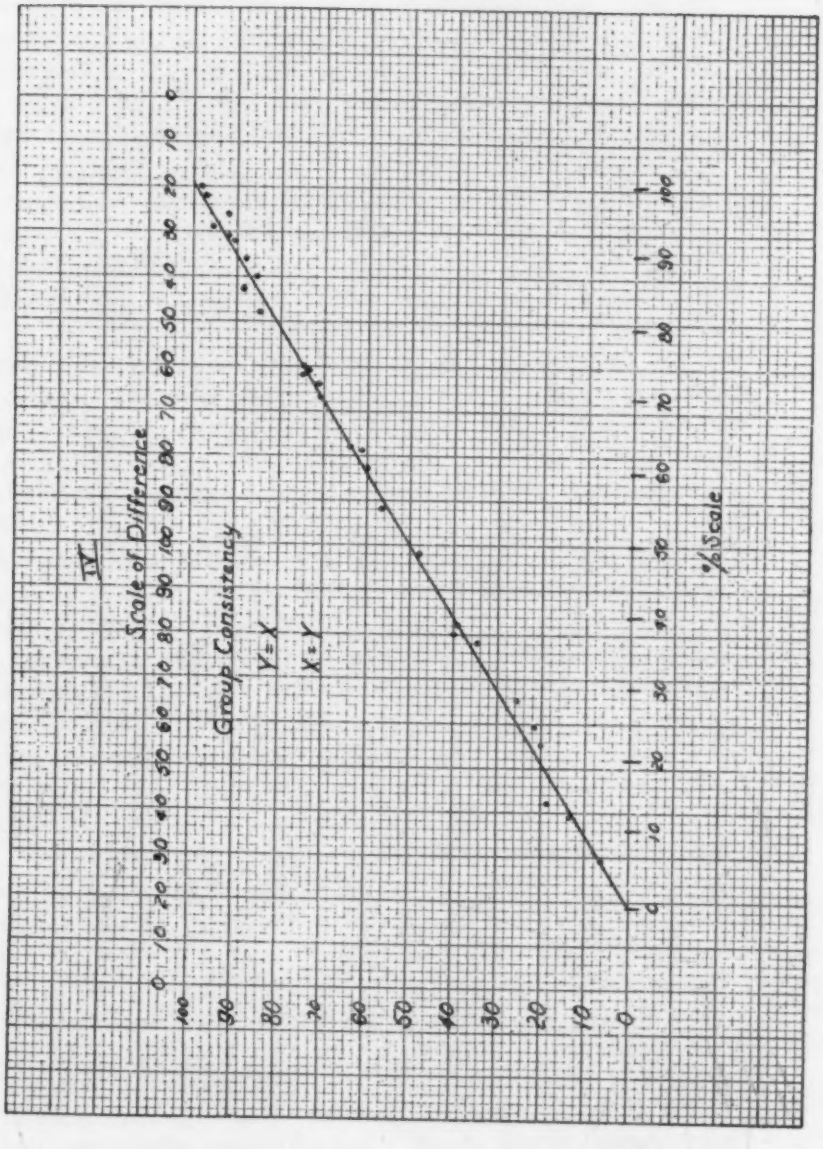
every other variable. The most fundamental relation, however, is that between the variables and the amount of stimulus difference in each of the experimental series.

It will be recalled that three different kinds of discrimination were tested, length, area, and weight. Since the shape of the curve may depend upon the kind of discrimination involved, the three forms were plotted separately, group consistency being the variable used. The results for length of line are shown in the first graph. It will be noticed that for values between .995 and .005, the best fit, as determined by eye, is a straight line extending from +19 to —19 on the scale. The curve for area is shown in the second graph. Again the curve of best fit, as determined by eye, is a straight line, extending from +19 to —19 on the scale. In other words, the curve for area is identical with the curve for distance. The curve for weights is given in graph three. Once again it is a straight line extending from +19 to —19 on the graph. Consequently, the three curves are identical in form.

For convenience of treatment, the scale of equal physical differences may be changed into a percentage scale. Let —19 = 0%, let +19 = 100%, let 100 = 50%, and the basis for the transformation is accomplished. This per cent scale will be used in the subsequent discussion.

Since the three curves are identical in form, they may be put together to form a single curve showing the relation between group consistency and multiples and submultiples of our unit difference. The result is shown in graph IV. Interpreted in terms of the percentage scale of difference, it shows a very strong tendency for *group consistency to vary directly with the difference.* The average distance of the dots from the line is .0143. In determining that $y$ varies directly with $x$, the following procedure was adopted. Each group consistency was expressed as a power or a root of the corresponding


Weights
III
100
90
80
70
60
50
40
30
20
10
0


IV
Scale of Difference
0 10 20 30 40 50 60 70 80 90 100
Group Consistency
Y=X
X=Y
100 90 80 70 60 50 40 30 20 10 0
% Scale
1 100 90 80 70 60 50 40 30 20 10 0

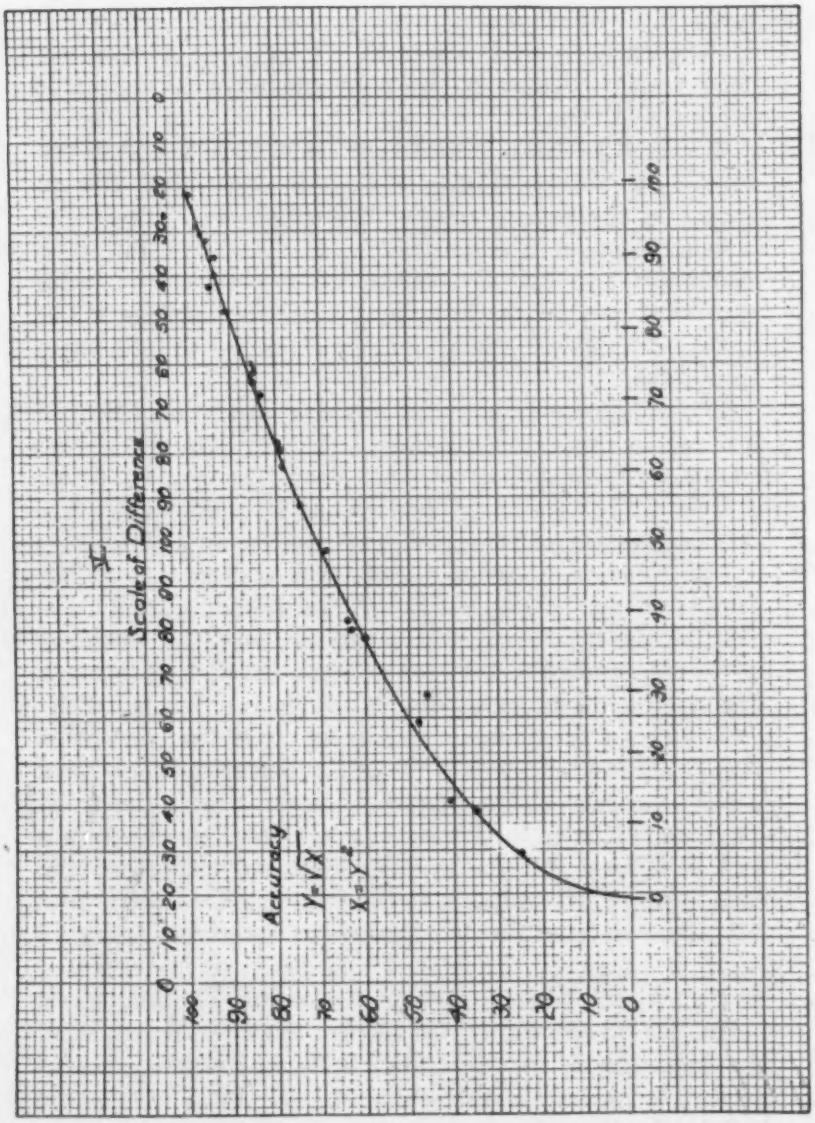
percentage value on the difference scale. The average value was 1.034 power with a very small dispersion. Consequently, a first power is a very close approximation, which, because of its convenience, was chosen to express the relation.

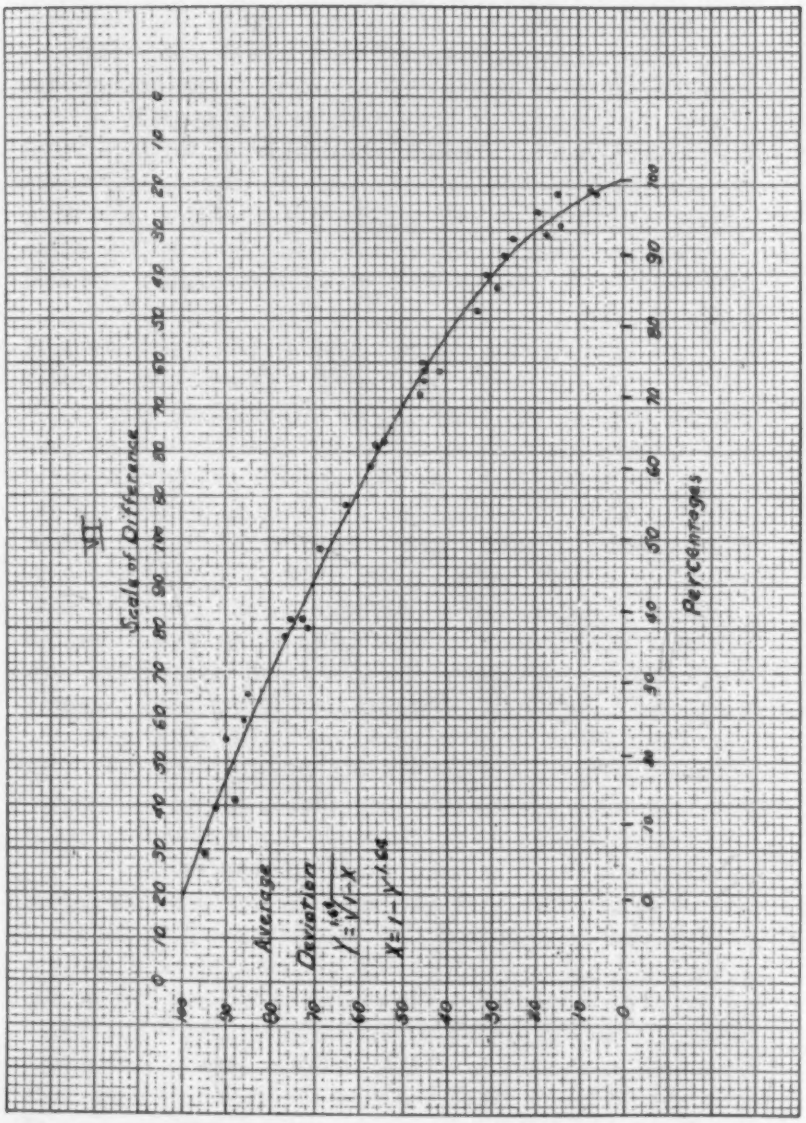
A second relation is that between accuracy and difference. As plotted in graph V, it appears that the accuracy is the square root of the difference. The mean displacement of the points from the curve is .0149. Moreover, the average relationship between accuracy and difference proves to be by actual computation a 1.98 root. The square root approximation is, consequently, very close.

The relationship between average deviation and difference is presented in graph VI. By the procedure adopted previously, it develops that the average deviation is the 1.64 root of one minus the difficulty. The average distance of the dots from the line is .0259.

The relation between standard deviation and difference is shown in graph VII. Again by computing roots and taking an average, it is found that the standard deviation equals the 2.1 root of one minus the difference. The average distance of the dots from the line is .0176.

In view of the evidence presented, it is our contention that we have obtained a scale of linear measurement which satisfies all of the requirements laid down by Titchener, Kuelpe, and Russell. We have provided an $x$ axis composed of equal units in the sense described and extending from minimum to maximum. The unit is that difference in length which can be correctly discriminated in 75 per cent of the cases. The scale is composed of multiples and submultiples of this unit. The maximum point is chosen arbitrarily as the smallest multiple of the unit that is discriminated with an accuracy represented by a correlation coefficient of .995. Past this point, an abrupt change in the direction of the curve takes place, for it increasingly approaches the horizontal.

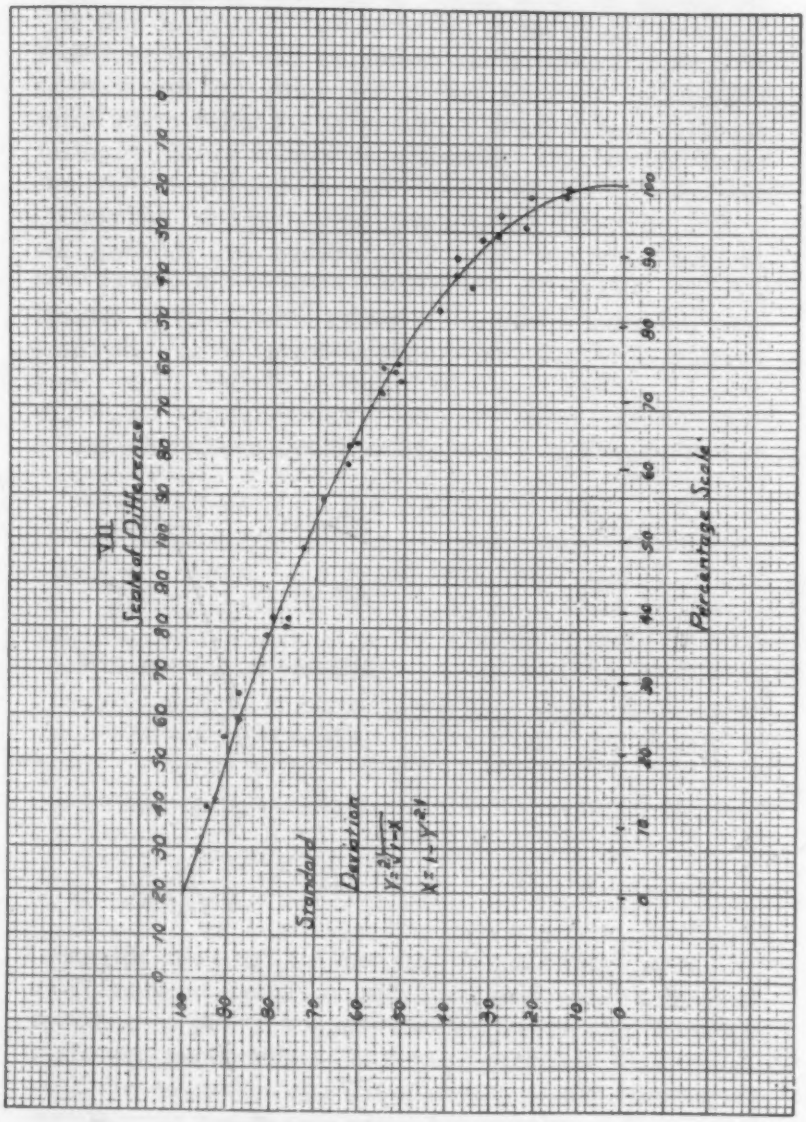

Scale of Diffraction
Accuracy
$Y=\sqrt{X}$
$X=Y^2$


Scale of Difference
Average Deviation
Percentages

Similarly, the minimum point is chosen arbitrarily as the largest submultiple of the unit that is discriminated with an accuracy represented by a correlation coefficient of .005. Beyond this point the curve turns sharply towards the horizontal. The scale is consequently a relative rather than an absolute physical scale, for it is based upon a least perceptible difference in extent, and many absolute differences will satisfy the condition, resulting in the same discriminable unit. At any rate the scale is based upon and derived from measurable physical units of extent.

The $y$ axis carries the correlation values representing the errors made in discriminating which of two or more objects possess the attribute to the greatest extent. It consequently depends upon judgments of more or less. These are found to be correlated with changes in the $x$ axis so that the coordinate $xy$ points form smooth curves. Consequently $y$ values can be translated into $x$ values, which are linear units. Such interpretation enables us to think of psychological findings in the more familiar physical terms, and permits us to compare results not only in terms of more or less, but also in terms of how much more or less. Quantitative expression becomes possible. When suitable arbitrary units are chosen and substituted for the original scale values, there is no reason why these units cannot be added, subtracted, multiplied, and divided. For the scale has a real zero point of just no ability to discriminate differences.

If it be admitted that a linear scale has been developed, it will also be admitted that a corresponding intensity scale has been attained, should discrimination of weight prove to be a function of intensity of kinaesthetic sensations. It coincides with the linear scale and is expressed in terms of the units of the distance scale. But attention should be called to the fact that the intensity scale has been derived from experiments on only 5 series of weights. These findings must be supplemented by ex-


Scale of Differences
Standard Deviation
Percentage Scale.

periments on other kinds of sensorial intensity before the scale can be taken too seriously. The great difficulty with such experimental procedure lies in the fact that changes of intensity are so frequently accomplished by simultaneous changes in quality of sensation or by other secondary criteria. Hence, such a scale would represent not intensity alone but a mixture of intensity and quality.

Not only has a linear scale been determined but mathematical equations have been worked out showing the relation between the $x$ values (scale of difference) and four different $y$ values. If group consistency, accuracy, average deviation and standard deviation are all functions of $x$, they should be functions of each other. And this is found to be the case. The equations which are to be regarded as *first approximations* are as follows:

$$GC = Acc^2 \qquad SD = {}^{2.1}\sqrt{1 - GC}$$

$$GC = 1 - SD^{2.1} \qquad SD = {}^{2.1}\sqrt{1 - Acc^2}$$

$$GC = 1 - AD^{1.64} \qquad SD = {}^{2.1}\sqrt{1 - AD^{1.64}}$$

$$Acc = \sqrt{GC} \qquad AD = {}^{1.64}\sqrt{1 - GC}$$

$$Acc = 1 - SD^{2.1} \qquad AD = {}^{1.64}\sqrt{1 - Acc^2}$$

$$Acc = 1 - AD^{1.64} \qquad AD = {}^{1.64}\sqrt{1 - SD^{2.1}}$$

A relationship which has not so far been discussed has very far-reaching implications. This is obtained from a comparison of self consistency and group consistency. Analysis of the data presented in this paper reveals a very strong tendency for *the self consistency correlations to be equal in value to the group consistency correlations.* Reference to the fundamental table will show that the

two values reveal a striking identity. But it also appears that the mean value of self consistency is .011 greater than the mean value of group consistency. This difference is characteristic, and, because of the very large number of cases involved, is presumably significant. The truth of the matter appears to be that the equality of the self consistency and the group consistency is an ideal to be striven for experimentally, but that this ideal is very infrequently realized in actual practice. A previous investigation* has shown that when the errors involved in the discrimination process are random in character, self consistency and group consistency are equal. As soon as systematic errors appear, however, self consistency is invariably greater than group consistency. The change in the relationship is due to the inflation of the self consistency, for group consistency and accuracy remain constant as the self consistency rises.

The constructive part of the paper may now be summarized. It has been found with the material studied that the errors involved in judgments of more or less are functionally related to the amount of difference discriminated. Thereby, a scale of equivalent physical units has been evolved in terms of which certain psychological processes can be measured. The measurement may be made in terms either of group consistency, accuracy, standard deviation, average deviation, or error, for each of these is not only a function of the difference, but each in turn is a function of every other one.

At this point, if not previously, the question is certain to be raised, of what value is a scale of this kind offered? In a general way, of course, any scientist will admit that quantitative knowledge is more useful than qualitative. Consequently, if the use of the scale can transform one single matter of opinion into a matter of fact, it has justified its existence. More specifically the writer can

* Adams, Journal of Social Psychology, February, 1930.

see several advantages accruing to psychology from the use of a quantitative scale.

In the first place since some four or five error variables have been related to physical differences, it is now possible to compare in quantitative terms the results of different researches in which the error has been expressed in different ways. To illustrate the point, the experiments of Snow, Hollingworth, and Adams on personality ratings may be cited. All three investigators employed the rank difference method. Snow used groups of 30 subjects, Hollingworth groups of 25, and Adams groups of 10. Hollingworth and Snow expressed their results in terms of average deviations, Adams presented his in terms of error. Referring these different variables to the scale of physical differences, Snow's results give a value equal to 42 percent on the scale, Hollingworth's a value of 42.4 percent, and Adams' a value of 41.6 percent. The average value is exactly 42 percent. This is in turn 8 percent below the just noticeable difference value of 50 percent, and defines with a high degree of precision the apparent difference between individuals with respect to the traits used in the investigations.

In the second place, the scale may be used to give numerical expression to differences which would otherwise be insignificant. Take the study made by Weber* on Moral Judgment in Female Delinquents. The Brogan list of 16 practices, ranging from sexual indulgence to dancing and smoking, was given to 138 inmates at the Geneva Reformatory in Nebraska. The rank difference method was used. Their results were compared with those obtained from a group of university women. Weber's conclusion, because the agreement between the two orders is +0.86, is, "The delinquency in these cases is not due to the absence of normal moral insight. Delinquents are approximately the equals of university

* Journal of Applied Psychology, Vol. 10, pages 89-91.

women." A study of the average deviations shows the average value for the university group to be 63 percent, for the Reformatory group 65 percent. Translated into terms of the scale of difference, these figures become 53 and 49. Both are very close to the unit value when the total list is taken into account. When we turn to the outstanding difference in the behavior pattern, namely, sexual indulgence, the average deviation for the university women is 49 percent, for the Reformatory women is 63.5 percent. These correspond to scale values of 70 and 53 respectively, the difference being much more obvious, 32 percent so, with the university women.

Another illustration may be taken from the experiments which were designed to measure the relative efficiency of advertising appeals. The average group consistency obtained from a considerable number of such investigations gives a value of approximately 20 percent on the scale of difference. This means that the differences between appeals are on the average so slight that it makes little difference which one is used.

A third use for the scale is found in the analysis of certain fundamental concepts, such as difficulty, which it permits us to make. Difficulty is a term which has crept into the literature largely in connection with tests and test items. By agreement items are of equal difficulty if they are reacted to correctly an equal number of times and the dispersions are identical, provided this number is neither 0 percent nor 100 percent of the possibilities. A comparison between the difficulty of two items which receive unequal numbers of correct responses has always been troublesome. Recourse to the area of the normal curve lying between this and that sigma has been the main method of evaluation.

Before going further, it is necessary to pause and ask what is meant by the word difficulty. At least three meanings will occur to anyone. It may describe the time required to do the task. But this meaning is ruled out

by the frequently heard remark that the job was long but not hard. Difficulty may be taken to mean the intensity of effort necessary to do a task. As difficulty increases, the available energy will not permit one to carry the task to completion. Consequently, the percentage of the task accomplished may be an index of the difficulty. Close to this in meaning is the one which relates difficulty to the number of errors made. The identification is complete if the amounts uncompleted be called errors. We shall define difficulty as measured by the percent of random errors. When this number is small, the task is easy; when the number is large, the task is hard. In the material derived from experiments on discrimination, it was found that the errors increased as the differences became smaller and smaller. It will be recalled that the group consistency varied directly with the amount of difference. Since group consistency is frequently used by the educational psychologists to measure difficulty, we may use the same term, and say that *difficulty varies inversely with group consistency.* In other words, group consistency measures ability directly, whereas 1— group consistency measures difficulty. Since all of the variables are interrelated, it is now possible to measure difficulty in terms of group consistency, self consistency, accuracy, average deviation, standard deviation, and error. Not only is it possible to do this but we can now compare different items in difficulty in terms of the scale of differences.

A fourth use is found in the critical examination of some of the statistical assumptions underlying the mathematical treatment of psychological data. Objectivity, reliability, and validity occur frequently as descriptive terms in the literature. Yet as applied to tests and scales the words have a wide variety of meaning. Our experiments enable us to check the statistical concepts against data derived from measurable material.

Objectivity has already been discussed. Experiments with objective scales show that self consistency and group consistency are equal. Similar experiments with subjective scales show that self consistency is always higher than group consistency. Here is a differentiating factor which associates subjectivity with systematic errors.

Reliability is a term which, like intelligence, may have had at some time in its career a definite meaning. That time, however, is long past. A return to first principles is obviously necessary. In our terminology, self consistency and group consistency may be substituted for reliability. Once this transfer has been made, it follows immediately that *reliability varies directly with the amount of the difference to be discriminated, and inversely with the difficulty of the discrimination.* This principle is definite, clear cut, and rational. It also permits the substitution of equivalents for it has been shown that when systematic errors are absent self consistency equals group consistency. Group consistency, when the test is valid, equals the square of the accuracy, the square root of 1 minus the standard deviation to the 2.1 power, and the square root of 1 minus the average deviation to the 1.64 power.

When systematic errors appear, however, self consistency gets out of step with the whole procession. It no longer equals anything but itself, yet the other relations all persist. It appears that any complete analysis of reliability must include a study of the effects of systematic errors upon self consistency, or objectivity. If systematic errors produce an inflation of self consistency and at the same time leave the group consistency values unchanged, then these errors introduce a spurious reliability into the test. For example, the group consistency of personality ratings is, as has been shown, +0.42. The self consistency is +0.76. Which of these values

gives the real reliability? Certainly it is not the .76; probably it is the .42. If one says that the reliability of personality ratings is .76, he must mean that there is a real reliability of .42, plus a spurious reliability derived from systematic errors which amounts to .34. This last figure is really an unreliability. With advertising appeals the self consistency is approximately .80, the group consistency about .20. Again, which figure represents the reliability? The answer is the same as before. Test 8 of the Army Alpha was changed to a 100 item true-false test. For the group used, the self consistency was .93, the group consistency .70. The conclusion is that reliability depends upon objectivity. Until the degree of objectivity of the test is known, the true reliability cannot be determined.

It is not unusual to see statements to the effect that a test is of little value unless it has a reliability of at least .90. A reliability as high as that can mean only that the test is very easy or that it is highly saturated with systematic errors. In view of the relations brought out in the earlier parts of the paper it would seem that the sensible procedure would be to determine reliability in terms of the other variables, group consistency, or standard deviation, rather than in terms of so unstable a measure as self consistency. For it makes no difference whether the order of items obtained from the concensus of opinion agrees with the true order of items, group consistency, average deviation and standard deviation will give the same values. On the other hand, only when the two orders agree, will accuracy and error values be at their optimum. Only in this case may we expect to find the accuracy equal to the square root of the group consistency.

Reliability is found to be affected by two other factors which deserve mention. In the first place it is found that reliability tends to increase as the number of test items increases. Presumably this is because the average

is better determined by the greater number of items. To ascertain the effect of the number of items upon reliability the Spearman-Brown prophesy formula has been proposed. This formula is

$$r_x = Nr/1+(N-1)r$$

In the second place, it has been found that reliability tends to increase as the results of more and more subjects are pooled to give a final order or value. Evidently the principle is the same as that involved in the first point and depends upon stabilizing the average. If so, the formula should apply to pools as well as to increased number of items. To test this point the material obtained from experiments 2 and 3, discrimination of lengths and areas, was pooled by 7's, 14's, 21's, 28's and 35's. The results are given in the table on the next page.

The two sets of figures run fairly close together, especially with supra-liminal differences. There is a slight tendency, however, for the obtained values to be greater than the theoretical, the mean difference being .0034. Disregarding signs so as to give the average divergence, the figure is .0078. The conclusion is that the prophesy formula will work as well for pools as for increased number of items. Since reliability is affected by both these factors, a statement should always be made of the number of items, or of the number of records pooled, for only when these facts are known can reliabilities be compared in any meaningful way.

We now turn our attention to validity. This term is used to designate a scale or test which measures what it is supposed to measure and nothing which it is not supposed to measure. In view of the last clause very few scales are entirely valid. In the case of the discrimination experiments in which one variable only was present, the validity is easily ascertainable. It can be identified with the accuracy, for the accuracy measures

| | | Lines | | Circles | |
|---|---|---|---|---|---|
| | | Obt. | Theor. | Obt. | Theor. |
| I | 1 | .856 | .856 | .918 | .918 |
| | 7 | .985 | .977 | .992 | .988 |
| | 14 | .994 | .988 | .998 | .994 |
| | 21 | .990 | .992 | .996 | .996 |
| | 28 | .995 | .995 | 1.000 | .998 |
| | 35 | 1.000 | .995 | 1.000 | .998 |
| II | 1 | .744 | .744 | .884 | .884 |
| | 7 | .973 | .953 | .984 | .981 |
| | 14 | .984 | .976 | .989 | .994 |
| | 21 | .987 | .984 | .993 | .995 |
| | 28 | .985 | .987 | .985 | .996 |
| | 35 | .980 | .990 | 1.000 | .997 |
| III | 1 | .614 | .614 | .845 | .845 |
| | 7 | .915 | .915 | .969 | .974 |
| | 14 | .956 | .957 | .984 | .988 |
| | 21 | .966 | .971 | 1.000 | .991 |
| | 28 | .985 | .979 | .995 | .994 |
| | 35 | .995 | .984 | .995 | .996 |
| IV | 1 | .479 | .479 | .745 | .745 |
| | 7 | .848 | .866 | .962 | .953 |
| | 14 | .960 | .929 | .980 | .976 |
| | 21 | .966 | .951 | .983 | .984 |
| | 28 | .990 | .962 | .990 | .987 |
| | 35 | .982 | .970 | .995 | .990 |
| V | 1 | .406 | .406 | .639 | .639 |
| | 7 | .791 | .826 | .929 | .926 |
| | 14 | .908 | .907 | .956 | .952 |
| | 21 | .932 | .935 | .991 | .975 |
| | 28 | .970 | .952 | .990 | .980 |
| | 35 | .990 | .960 | .985 | .984 |

the ability of the subjects to discriminate. As in the case of reliability, accuracy tends to vary with the size of the difference. Reference to the scale shows that validity varies as the square root of the difference. This again is simply another way of saying that validity is the square root of the true reliability. The same statement has been made by May.*

Unless the test is known to be valid the validity can be determined only from the accuracy, or from the error, of the variables used in this paper. For these are the only values which are affected by the relationship between the order determined by consensus of opinion and the true order. Self consistency, group consistency, average deviation, and standard deviation will give no index of the validity of the test. The relation between group consistency and accuracy may, however, be an index of validity. For when the group consistency is not the square of the accuracy, the order determined by concensus of opinion will not agree with the objective order, and the test is not valid.

Whenever the order determined by consensus of opinion and the true order do not agree, it is because one or more of the items is reacted to or discriminated incorrectly in more than 50 percent of the cases. The error becomes unduly large and the accuracy correlation correspondingly small. Now in what situation may we expect to find an error greater than 50 percent? Since chance would give in the long run 50 percent of errors, some factor other than chance would have to be operative. Random errors, which are due to chance, may be eliminated from any study of validity. Two other kinds of error are possible, the systematic and the constant. The systematic errors have already been dealt with and it has been shown that they affect reliability but do not

* Studies in Deceit, Book 2, pages 122 f.

necessarily affect accuracy. It is probable that they never do, in any serious way, affect accuracy, though they make correction of errors a slower process.

As far as errors are concerned it is probable, therefore, that constant errors are responsible for lack of validity. To illustrate we may cite another discrimination experiment. This was a repetition of experiment 3, only this time the circles instead of being drawn on separate cards were all 10 for each series drawn on the same sheet of paper. The result was that in none of the five series did the experimental order of size agree with the actual order of size. Illusions resulting from the juxtaposition of the different circles emphasizing confluxion and contrast, together with the fact that there was a persistent tendency to overestimate those at the top of the page and underestimate those at the bottom, all cooperated to produce constant errors. In all cases the reliability was affected to only a very slight extent but the validity was very much reduced. This is shown by the correlations between the experimental order and the true order, which ranged from +.96 to +.56. Since constant errors are the only kind which can increase the errors to more than 50 percent, and since errors of 50 percent or more are necessary to produce a non-valid scale, it seems safe to conclude that constant errors are the cause of the validity of scales.

There appears to be a confusion between validity and precision. Yet a moment's thought will show that a scale may be valid but not precise, or precise but not valid. As has been shown validity varies as the square root of the difference to be discriminated. Precision is increased by pooling the results of many subjects, thereby becoming a group phenomenon. Validity, on the other hand, is determined by an average, not by a summation.

It must be understood that the foregoing remarks

apply only to the results obtained from the discrimination experiments described in the paper. By implication they may be included to apply also to intelligence tests when these are tabulated item by item. When this method has been used self consistency, group consistency and accuracy are interrelated in the ways already described.

# AN EXPERIMENTAL STUDY OF VISUAL APPREHENSION IN READING*

OSCAR F. LITTERER
*University of Minnesota*

Although it seems obvious that visual apprehension is an important determinant of reading performance, no study employing adult subjects has been published on its relationship to reading achievement when reading is measured by standardized tests. Gates (2, 3) has made two studies on elementary school pupils and Sister Mary of the Visitation (5) has performed a similar one. This study is an attempt to discover the approximately true relationship for adult subjects between visual apprehension and achievement on reading and intelligence tests. This is done by the correlation method; and since the reliability of the tests affects directly the relationships studied, it is taken into account in the interpretation of the data.

Perception tests composed of letters, digits, words and sentences, of twenty series each, were developed to measure visual apprehension. The material of each series was hand printed on a strip of white cardboard 28 inches by 11 inches. The letters and the digits were printed in one line, the words in two lines, and the sentences in three lines. Thorndike's (6) one thousand most frequently used word list was utilized to select the words. The sentences of fifteen words each were selected from elementary school readers. In presentation of stimuli provision was made to equate practice effects.

A simple exposure apparatus was employed to expose the printed cardboards for an interval of three seconds. A preliminary test of six sentences was given to acquaint

* Part of a reading investigation done in the Psychological Laboratory, University of Minnesota, under the direction of Professor Miles A. Tinker.

the subjects with the procedure. The subjects were instructed to reproduce in the precise order at the end of the interval the content apprehended during the exposure. The average visual apprehension span was computed on the basis of a credit of one for each unit reproduced correctly and in the right place in the series and a credit of one-half for each unit reproduced correctly but in a changed position in the series. The subjects were tested in groups of from 30 to 40 students.

Reading performance was measured by rate and comprehension tests. Forms A and B of the Chapman-Cook Speed of Reading Test, and Form A of the Minnesota Speed of Reading Test were employed as measures of the rate of reading. The two tests differ mainly in the difficulty of content.

The Minnesota Reading Examination and two parts of the Iowa Silent Reading Examination were used to secure a score of the comprehension. The Minnesota Examination consists of a vocabulary and a paragraph reading test which were scored independently and are treated as separate tests in the data. The paragraph comprehension and paragraph organization tests of the Iowa Examination are also presented as separate tests in the study.

Intelligence was measured by the college ability test used at the University of Minnesota and by Form A of the Army Alpha Examination.

The subjects in this study were sophomore college students who registered for the laboratory psychology course. It was found impossible to secure a score from all subjects on each test and, as a result, the number of subjects on the various comparisons range between 122 and 176.

The perception tests measure visual apprehension with a high degree of consistency as is indicated by the reliability coefficients presented in Table I. The coeffi-

cients, after the Spearman-Brown formula was applied, range from .841 to .929. The reliability correlations were computed by the odd-even method—unless otherwise stated—and were used to correct the inter-correlations for attenuation.

All reading measures except the Minnesota Paragraph Reading Test showed high internal consistency. The reliability of .611 for paragraph reading is adequate for group comparisons only.

The inter-correlations of the perception tests in Table II indicate a trend which is similar to the one observed

TABLE I

*Reliability Coefficients*

| Tests | Raw r | Raw P. E. | S—B Prophesy Formula |
|---|---|---|---|
| Visual Apprehension Tests: | | | |
| Letters | .726 | ± .02 | .841 ± .01 |
| Digits | .838 | ± .01 | .911 ± .01 |
| Words | .851 | ± .01 | .919 ± .01 |
| Sentences | .868 | ± .01 | .929 ± .00 |
| Chapman-Cook Speed of Reading Test | .896 | ± .01 | .945 ± .00 |
| Minnesota Speed of Reading Test | .860* | | .860* |
| Minnesota Reading Examination: | | | |
| Vocabulary Test | .902 | ± .01 | .948 ± .00 |
| Paragraph Reading Test | .440 | ± .04 | .611 ± .03 |
| Iowa Silent Reading Examination: | | | |
| Paragraph Comprehension Test | .815 | ± .01 | .898 ± .01 |
| Paragraph Organization Test | .888 | ± .01 | .940 ± .00 |
| Army Alpha Examination | .897 | ± .01 | .945 ± .00 |
| College Ability Examination | .850* | | .850* |

* A reliability coefficient was taken from Eurich's (1) investigation for the Minnesota Speed of Reading Test and the one cited for the College Ability Examination was secured from Paterson and Drake's (7) report. These two coefficients and the one for the Chapman-Cook Speed of Reading Test were derived by correlating scores on two forms of the tests and, hence, are not strictly comparable to the odd-even coefficients in the rest of the table.

by Gates (3) and Sister Mary of the Visitation (5). Both investigators found that, "tests of the same content

TABLE II

*The Inter-correlations of the Visual Apprehension Tests*

| Tests | Raw r | Raw P. E. | Atten. r | Atten. P. E. |
|---|---|---|---|---|
| Letters vs. Digits | .719 | ± .02 | .821 | ± .01 |
| Letters vs. Words | .659 | ± .03 | .749 | ± .02 |
| Letters vs. Sentences | .388 | ± .04 | .439 | ± .04 |
| Digits vs. Words | .555 | ± .03 | .607 | ± .03 |
| Digits vs. Sentences | .327 | ± .04 | .355 | ± .04 |
| Words vs. Sentences | .705 | ± .02 | .763 | ± .02 |

but different in form correlated higher than tests the same in form and different in content." The relative importance of content and form was not investigated but it is evident that the nature of the content has a decided influence on the inter-correlations. Letters and digits show the lowest relationship with sentences and, on the other hand, letters with digits, and words with sentences correlated the highest. The range of the coefficients is large. It extends from .327 to .719 and when corrected for attenuation from .355 to .821. The average inter-correlation is .558, and corrected, .622, which is significantly higher than those reported by the above mentioned investigators, and may be attributed to the greater similarity in the content.

The nature of perception has received some attention. Gates (3) concluded that perception is highly specialized, because only tests composed of words correlated substantially with reading and spelling. Sister Mary of the Visitation (5), however, came to the opposite conclusion. By a general inspection of the inter-correlations in her study, she observed some evidence in favor of a group factor as described by Spearman in "The Abilities of Man" or, in other words, the existence of a general per-

ceptual ability. The partial correlation technique was applied to the data at hand with the hope of revealing some evidence pertaining to the problem. Sentences correlated the lowest with digits and highest with words. A partial correlation was obtained between digits and sentences with words held constant and the $r$ was found to be –.210. Similarly the correlation between digits and sentences with letters held constant was found to be –.009. Apparently there is no common factor running through both digits and sentences. Two partial correlations were, furthermore, computed between variables of similar materials. Words were correlated with sentences and digits held constant. The coefficient is .739 which is only .024 lower than the original $r$ between words and sentences and suggests that digits have practically no influence on the other two variables. The other partial correlation was computed between letters and digits with sentences held constant. This coefficient (.794) is, likewise, only .027 lower than the original correlation between letters and digits. The results suggest that there is no general factor common to the meaningless and the meaningful materials used in this study.

Inter-correlations of the reading tests are given in Table III. The paragraph reading test of the Minnesota Reading Examination correlated low with all other measures of reading. This may be partly a function of the low reliability of the Minnesota test. All the remaining inter-correlations, however, fall between .463 and .606 which indicates a fair relationship among the various kinds of reading.

The magnitude of the correlations between the mental tests and reading tests shows an appreciable factor common to the two kinds of performance. The average correlation, corrected for attenuation, is .619 with a range from .449 to .808 (Table IV). Rate and comprehension in reading, as measured by the tests employed in this study, depend upon approximately the same de-

gree of mental ability. The average correlation between intelligence and reading was found to be .545 which is practically the same as those obtained by Gates (2, 3) between the Stanford-Binet and reading. The vocabu-

TABLE III

*The Inter-correlations of the Reading Tests*

| Tests | Raw r | Raw P. E. | Atten. r | Atten. P. E. |
|---|---|---|---|---|
| Chapman-Cook vs. Vocabulary Test (M. R. E.)* | .606 | ± .03 | .640 | ± .03 |
| Chapman-Cook vs. Paragraph Reading Test (M. R. E.) | .285 | ± .05 | .375 | ± .05 |
| Chapman-Cook vs. Paragraph Comprehension (I. S. R. E.) | .640 | ± .03 | .694 | ± .02 |
| Chapman-Cook vs. Paragraph Organization (I. S. R. E.) | .583 | ± .03 | .618 | ± .03 |
| Chapman-Cook vs. Minnesota Speed Test | .569 | ± .03 | .631 | ± .03 |
| Minnesota Speed Test vs. Vocabulary Test (M. R. E.) | .515 | ± .04 | .570 | ± .03 |
| Minnesota Speed Test vs. Paragraph Reading (M. R. E.) | .045 | ± .05 | .062 | ± .05 |
| Minnesota Speed Test vs. Paragraph Comprehension | .514 | ± .04 | .585 | ± .03 |
| Minnesota Speed Test vs. Paragraph Organization | .501 | ± .04 | .557 | ± .04 |
| Vocabulary Test vs. Paragraph Reading Test | .463 | ± .04 | .608 | ± .03 |
| Vocabulary Test vs. Paragraph Comprehension Test | .579 | ± .03 | .627 | ± .03 |
| Vocabulary Test vs. Paragraph Organization Test | .546 | ± .04 | .579 | ± .04 |
| Paragraph Reading Test vs. Paragraph Comprehension Test | .276 | ± .05 | .372 | ± .05 |
| Paragraph Reading Test vs. Paragraph Organization Test | .324 | ± .05 | .428 | ± .04 |
| Paragraph Comprehension vs. Paragraph Organization | .546 | ± .04 | .594 | ± .03 |

* M. R. E. designates the Minnesota Reading Examination, and I. S. R. E. designates the Iowa Silent Reading Examination.

lary test, however, shows a somewhat higher relationship to the same intelligence tests (.763 and .808).

The visual apprehension of meaningless material is not associated with reading performance of adult subjects to any appreciable degree. Scores for letters (Table V) correlate to a slight degree with the reading tests. The average coefficient, corrected for attenuation, is .283 with a range from .165 to .501. The apprehension of digits, however, has practically no association with reading. Meaningful material, as exemplified by words and sentences, correlates somewhat higher than either digits or letters with reading performance. The

TABLE IV

*The Correlations of the Mental Tests with the Reading Tests*

| Tests | Raw r | Raw P. E. | Atten. r | Atten. P. E. |
|---|---|---|---|---|
| Army Alpha vs. Chapman-Cook Speed Test | .586 | ± .03 | .620 | ± .03 |
| Army Alpha vs. Minnesota Speed Test | .405 | ± .04 | .449 | ± .04 |
| Army Alpha vs. Paragraph Comprehension Test (I. S. R. E.) | .558 | ± .04 | .605 | ± .03 |
| Army Alpha vs. Paragraph Organization (I. S. R. E.) | .552 | ± .04 | .585 | ± .03 |
| Army Alpha vs. Vocabulary Test (M. R. E.) | .722 | ± .02 | .763 | ± .02 |
| Army Alpha vs. Paragraph Reading Test (M. R. E.) | .552 | ± .04 | .727 | ± .02 |
| College Ability vs. Chapman-Cook Speed Test | .521 | ± .04 | .581 | ± .03 |
| College Ability vs. Minnesota Speed Test | .499 | ± .04 | .584 | ± .03 |
| College Ability vs. Paragraph Comprehension (I. S. R. E.) | .526 | ± .04 | .602 | ± .03 |
| College Ability vs. Paragraph Organization (I. S. R. E.) | .474 | ± .04 | .530 | ± .04 |
| College Ability vs. Vocabulary Test (M. R. E.) | .725 | ± .02 | .808 | ± .02 |
| College Ability vs. Paragraph Reading Test (M. R. E.) | .420 | ± .04 | .583 | ± .03 |

TABLE V

*The Correlations of the Visual Apprehension Tests with the Reading Tests*

| Tests | Raw | | Atten. | |
|---|---|---|---|---|
| | r | P. E. | r | P. E. |
| Letters vs. Chapman-Cook Speed Test | .201 | ± .05 | .225 | ± .05 |
| Letters vs. Minnesota Speed Test | .177 | ± .05 | .208 | ± .05 |
| Letters vs. Paragraph Comprehension Test (I. S. R. E.) | .232 | ± .05 | .266 | ± .05 |
| Letters vs. Paragraph Organization Test (I. S. R. E.) | .147 | ± .05 | .165 | ± .05 |
| Letters vs. Vocabulary Test (M. R. E.) | .300 | ± .05 | .336 | ± .05 |
| Letters vs. Paragraph Reading Test (M. R. E.) | .359 | ± .05 | .501 | ± .04 |
| Digits vs. Chapman-Cook Speed Test | .022 | ± .05 | .023 | ± .05 |
| Digits vs. Minnesota Speed Test | .128 | ± .05 | .144 | ± .05 |
| Digits vs. Paragraph Comprehension Test (I. S. R. E.) | .159 | ± .05 | .175 | ± .05 |
| Digits vs. Paragraph Organization Test (I. S. R. E.) | .094 | ± .05 | .101 | ± .05 |
| Digits vs. Vocabulary Test (M. R. E.) | .177 | ± .05 | .190 | ± .05 |
| Digits vs. Paragraph Reading Test (M. R. E.) | .208 | ± .05 | .278 | ± .05 |
| Words vs. Chapman-Cook Speed Test | .421 | ± .04 | .452 | ± .04 |
| Words vs. Minnesota Speed Test | .239 | ± .05 | .269 | ± .05 |
| Words vs. Paragraph Comprehension Test (I. S. R. E.) | .387 | ± .04 | .426 | ± .04 |
| Words vs. Paragraph Organization Test (I. S. R. E.) | .242 | ± .05 | .260 | ± .05 |
| Words vs. Vocabulary Test (M. R. E.) | .313 | ± .05 | .335 | ± .05 |
| Words vs. Paragraph Reading Test (M. R. E.) | .332 | ± .05 | .443 | ± .04 |
| Sentences vs. Chapman-Cook Speed Test | .438 | ± .04 | .467 | ± .04 |
| Sentences vs. Minnesota Speed Test | .365 | ± .04 | .408 | ± .04 |
| Sentences vs. Parapragh Comprehension Test (I. S. R. E.) | .490 | ± .04 | .536 | ± .04 |
| Sentences vs. Paragraph Organization Test (I. S. R. E.) | .377 | ± .05 | .403 | ± .04 |
| Sentences vs. Vocabulary Test (M. R. E.) | .379 | ± .05 | .404 | ± .04 |
| Sentences vs. Paragraph Reading Test (M. R. E.) | .288 | ± .05 | .382 | ± .05 |

apprehension of sentences tends to be more closely associated with reading than words. The average corrected coefficient is .433 with a range from .382 to .536. These correlations suggest that the apprehension of either words or sentences is only a secondary factor in reading of adult subjects.

The visual apprehension material yields corrected coefficients which range from .200 to .467 when correlated with the mental examinations (Table VI). The length

TABLE VI

*The Correlations of the Mental Tests with the Visual Apprehension Tests*

| Tests | Raw r | Raw P. E. | Atten. r | Atten. P. E. |
|---|---|---|---|---|
| Army Alpha vs. Letters | .398 | ± .05 | .446 | ± .04 |
| Army Alpha vs. Digits | .316 | ± .05 | .340 | ± .05 |
| Army Alpha vs. Words | .370 | ± .05 | .397 | ± .05 |
| Army Alpha vs. Sentences | .438 | ± .04 | .467 | ± .04 |
| College Ability vs. Letters | .263 | ± .05 | .311 | ± .05 |
| College Ability vs. Digits | .176 | ± .05 | .200 | ± .05 |
| College Ability vs. Words | .248 | ± .05 | .280 | ± .05 |
| College Ability vs. Sentences | .378 | ± .05 | .425 | ± .04 |

of the range suggests that intelligence, as measured by these tests, shows a slight and fairly constant relation to visual apprehension regardless of material.

In general visual apprehension tends to be slightly more associated with intelligence than reading. This is contrary to the findings on elementary pupils (3, 5). Gates (3) used the partial correlation technique to determine the correlation between perception and reading with the influence of intelligence eliminated. The partial correlations were reduced by only a slight amount and these coefficients still represented a relationship higher than was secured for perception and intelligence. In this study the letters and the sentences were selected of the meaningless and meaningful materials respectively

to compute partial correlations because they show the closest association with reading. Visual apprehension of letters in nonsense series shows little or no relationship to reading when intelligence is held constant. Letters yield an average correlation of .015 (range: –.131 to .288) and .127 (range: .001 to .415) with the reading tests when the Army Alpha and the College Ability Examinations respectively are held constant. Sentences, however, show slightly more correspondence to reading when intelligence is eliminated: the average partial correlations are .200 (range: .070 to .361) and .239 (range: .114 to .389) respectively. Apparently, the relation of visual apprehension of letters to reading performance is largely a factor of intelligence as measured by these verbal group examinations while the relation of sentence apprehension to reading is only partly due to intelligence.

## Summary and Conclusion

1. In this study, an attempt was made to discover the approximately true relationship that exists among visual apprehension and achievement on reading and on intelligence tests when college students are used as subjects. The relative importance of visual apprehension and of intelligence in reading performance was also investigated.

2. The visual apprehension of meaningless materials is not significantly associated with reading. The letters and digits have average corrected coefficients of .283 and .151 with the reading tests respectively. The meaningful material has a low but statistically significant relationship to reading. The average corrected coefficients obtained for words and sentences with the reading tests are .322 and .433 respectively. In general, visual apprehension does not show a high correspondence with scores on reading tests.

3. Intelligence shows a low but rather constant correspondence with the visual apprehension materials. The range of the corrected coefficients is from .200 to .467. The perception tests used in this study, apparently, involve a fairly constant amount of intelligence.

4. The visual apprehension of meaningless and meaningful material is slightly more closely related to intelligence than to reading. This is contrary to the findings on elementary school children. The discrepancy may be due either to the different forms of the perception tests employed or to the difference in age of subjects.

## BIBLIOGRAPHY

1. Eurich, A. C.: *The Reading Abilities of College Students.* University of Minn. Press, Minneapolis, 1931, 55-56.

2. Gates, A. I.: *The Psychology of Reading and Spelling.* Teachers College, Columbia University. Contributions to Education, No. 129. Teachers College, Columbia University, New York City, 1922.

3. Gates, A. I.: *A Study of the Role of Visual Perception, Intelligence and Certain Associative Processes in Reading and Spelling.* J. Educ. Psychol., 1926, 17, 433-445.

4. Paterson, D. G. and Drake, L. E.: *Reliability Analysis of 1929 Edition of College Ability Tests Prepared for Minnesota Association of Colleges.* (Not published.)

5. Sister Mary of the Visitation: *Visual Perception in Reading and Spelling: A Statistical Analysis.* Educ. Res. Bull., The Catholic University of America, Vol. IV, Jan., 1929, No. 1.

6. Thorndike, E. L.: *The Teacher's Word Book.* Teachers College, Columbia University, New York City, 1921.

# AN ATTEMPT TO PREDICT THE SUCCESS OF STUDENT NURSES BY THE USE OF A BATTERY OF TESTS*

JESSE BATLEY RHINEHART
*University of Pittsburgh*

## I. The Problem

The problem of selection of students is of vital economic significance in nursing education. In the majority of schools the elimination of undesirables is done during a six-month probation period during which the quality of work determines rejection or final admission. Much time and money are wasted by nursing schools in this trial-and-error procedure.

Several methods of selection have been suggested as means of decreasing the percentage of dismissals during the probation period. One method attempts a solution by increasing the entrance requirement from "one year of high school work" to "graduation from a four-year high school." Another goes further, suggesting the acceptance of only those high school graduates whose grades rank them in the upper half of the class. Advocates of such plans assume that by requiring a longer period of academic preparation and by raising scholastic requirements fewer failures will be found in probation groups. A third method suggests the administration of intelligence tests to applicants and recommends immediate rejection of those scoring low. Practically all leaders in nursing education agree that a predictive measure of personality and attitude would be of inestimable value in selection of students.

This study represents an attempt to determine the value of a battery of tests in predicting success in nurs-

* Digest of unpublished thesis written under the direction of Dr. Florence M. Teagarden at the University of Pittsburgh in University Library.

ing by comparing the scores made on such tests by a group of student nurses with their grades in *academic* and *practical work*. The tests used in this study include a personality inventory, an individual intelligence test, a group intelligence test, a social intelligence test, and a nursing aptitude test.

## II. Previous Studies

Actual experimental investigations for the purpose of finding prognostic tests that would effectively serve to supplement the usual criteria for admission to schools of nursing are few. Those investigations which have been made have for the most part employed intelligence tests almost exclusively.

McPhail (5) administered the Brown University Test and the Army Alpha Test to a group of student nurses at the Rhode Island Hospital. The predictive value of scores was studied. McPhail concluded that the psychological scores are of predictive significance among probationers and also among juniors so far as the academic grades are concerned. This conclusion is not particularly enlightening since we already know that there is a fair positive correlation between intelligence test scores and academic grades.

Earle (1), who administered the Army Alpha Test to a group of student nurses, says: "One might well consider eliminating those making an Alpha grade of below 75, provided always that the results of this test were corroborated by other subjective and objective measures as reliable." Earle also concludes that personality and emotional traits are as valuable if not more so than general intelligence.

Metcalfe (4) reports a study by a committee of the Nursing Education League in which the records of 439 student nurses were studied. The general conclusion reached in this study is that the Army Alpha Test does

not measure the total nursing ability of an individual; it gives an indication of general intelligence but does not test many qualities that good nurses have—qualities that should be considered or required in selecting students for nursing.

South and Clark (6) after administering a number of group intelligence tests, educational tests, and both forms of the Colgate Personality Inventory "feel that the following traits are representative of the nursing personality: intellectual, scholastic, emotional stability, social attitude."

Faber (2) cites some reasons favorable to the use of intelligence tests in nursing schools. She concludes that such testing makes possible the selection of a higher type of woman, is useful in cases of maladjustment, lessens delay in determining the acceptability of the student, and finally, that an intelligence test might be the determining factor in cases of doubtful acceptability.

In agreement with the conclusions of McPhail and of Earl, Hyman and Dreyfuss (3) state that differences in intelligence do not have so much weight in determining future success in nursing as does the possession of other characteristics such as attitude and personality.

Young (8) in a study of a group of student nurses at Indiana University found: "The students rating highest on the intelligence test made distinctly the highest average in both their theoretical studies and their efficiency records on practical work."

The studies cited here seem to regard favorably the prognostic value of intelligence tests in schools of nursing but practically all suggest that other tests measuring emotional and personality traits would be of value in predicting success. Faber (2) suggests that social intelligence is particularly essential. Others interested in the problem of selection of students for schools of nursing have suggested the use of aptitude tests for the purpose of prognostication.

This investigation is unique in that:

(a) It attempts to determine experimentally the predictive value of a social intelligence test (the Moss Social Intelligence Test) and a personality test (the Bernreuter Personality Inventory).

(b) It attempts to determine the predictive value of an aptitude test (the Moss Nursing Aptitude Test).

(c) It is the first study to evaluate the Stanford-Binet IQ for the purpose of predicting success in nursing.

(d) It attempts to determine whether higher academic entrance requirements have operated to select students superior to those admitted under less rigid conditions.

## III. Method

The present study, which was made in a Pittsburgh training school for nurses, employed a battery of tests consisting of: the Stanford-Revision of the Binet-Simon Intelligence Test, the American Council on Education Group Intelligence Test, the Moss Social Intelligence Test, the Moss Nursing Aptitude Test, and the Bernreuter Personality Inventory. The entire battery of tests was given to 48 first-year student nurses, all of whom were high school graduates, ranking in the upper half of their classes. The Moss Nursing Aptitude Test was also administered to 46 third-year students, half of whom were high school graduates. The students making up the remaining half of this class varied in high school preparation from one to three years.

The statistical treatment consisted primarily of correlating scores made on the tests with grades in the *theory* and the *practice* of nursing. The grade in *theory* used here is the average of the first semester's theoretical work, which consists of courses such as anatomy, bacteriology, ethics, dietetics, and chemistry. The theoretical work is, of course, practically all of an academic nature. The grade in *practice* is an estimation of the student's performance in that part of the training course

consisting of actual work or "floor duty." The student spends a certain number of hours each day "on duty," and during this time she actually cares for patients. The grade in this practical work is recorded on the Efficiency Record, a form required by the state of Pennsylvania. On the Efficiency Record the student is graded in percentages on such items as acceptability to patient, reliability, courtesy, cheerfulness, dignity, and sympathy. The head nurse does the grading. The most severe criticism of this measure of practical work lies in its lack of objectivity in scoring. This lack of objectivity makes the Efficiency Record an inadequate and inaccurate measure of the student's worth in practical nursing. The degree of relationship between the scores on the tests administered and the grades in practical work might have been greater if the practice grade had been a composite score, obtained from the scorings of several persons, and if the rating scale used had been a reliable one.

## IV. Results

### 1. *Intelligence as measured by the Binet-Simon Test (Stanford Revision).*

In the controversy over selection of students for schools of nursing, attempts have been made to set an arbitrary minimum IQ as an additional entrance requisite. Dr. Q. A. W. Rohrbach[1] suggests that no student whose IQ is less than 110 should be admitted into the training school since, he says, individuals rating below this minimum will not be able to complete successfully the training course.

The Stanford Revision of the Binet-Simon Intelligence Test was administered to 47 first-year students by the writer. In calculating the IQ year 16, as suggested by

[1] In an address to the Pennsylvania League of Nursing Education given in October 1931 at Pittsburgh, Pa.

Terman (7), was used as the divisor. The range of IQ's was found to be 71 to 122.

A study of the results below indicates that individuals below the adult norm in intelligence *are* capable of doing satisfactory work in the training school in which this investigation was made. Terman places normal intelligence between the IQ's of 90 and 110. With this in mind, the first-year students were divided into three groups, the superior consisting of IQ's above 110, the normal being made up of IQ's between 90 and 110, and the subnormal group representing those IQ's below 90. The distribution of IQ's places 13 per cent in the superior group, 55 per cent in the normal group, and 32 per cent in the subnormal group. Table I lists the average grades in theoretical and in practical nursing for these three groups.

TABLE I

*Average Grades for Groups which Include the Superior, Normal, and Subnormal IQ's*

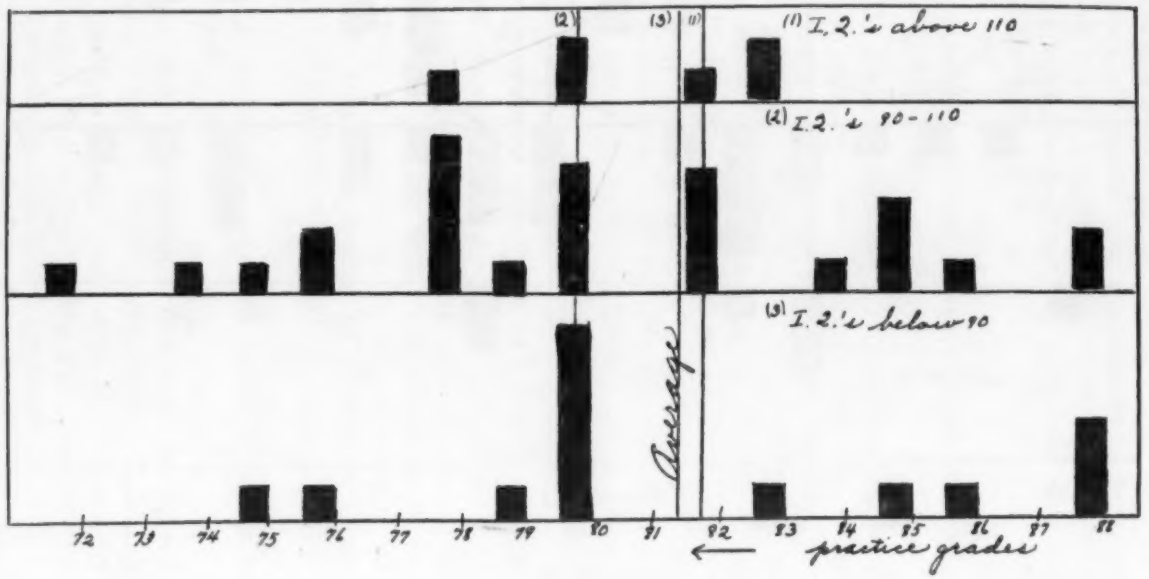
| Groups | Average Theory Grade | Average Practice Grade |
|---|---|---|
| Superior 13% | 87.7 | 81.8 |
| Normal 55% | 83.8 | 79.9 |
| Subnormal 32% | 83.2 | 81.6 |

Inspection of Table I shows that the upper 13 per cent excels the lower 32 per cent 4.5 points in theory and .2 in practice. The middle 55 per cent scores only 3.9 less in theory and 1.9 less in practice than does the upper 13 per cent.

In Figures 1 and 2 the same results are shown graphically. Note that the middle and lower groups spread over practically the same range in theoretical grades. Note also that the lower group actually excels the middle group in practical work. An examination of this graph shows that the adoption of an IQ of 110 as an entrance requisite would exclude many students who are capable

FIGURE 1

*Distribution of IQ's according to practice grades. Each perpendicular line represents the average grade for that group.*



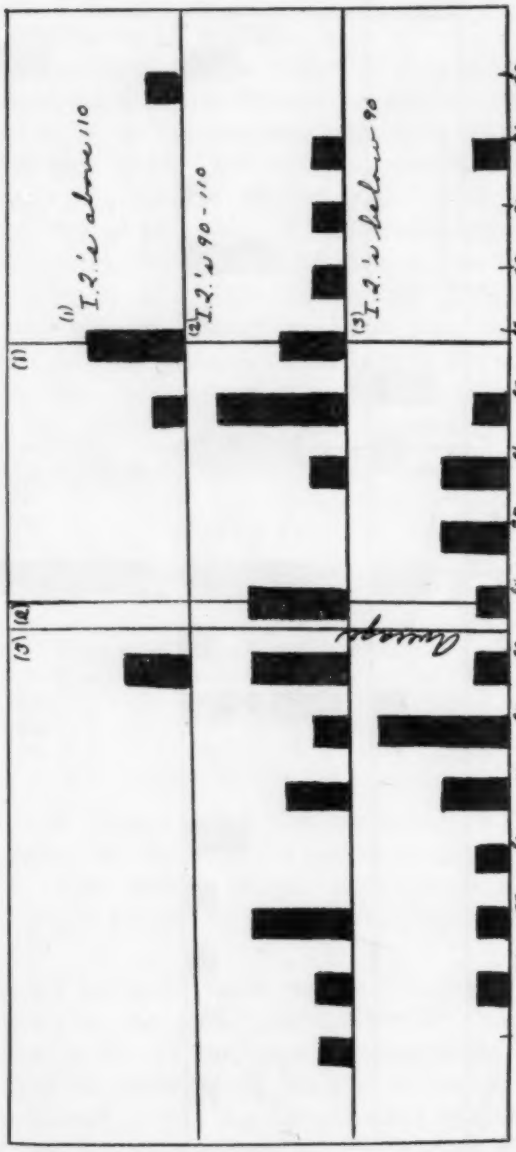



FIGURE 2

*Distribution of IQ's according to theory grades. Each perpendicular line represents the average grade for that group.*

of performing satisfactorily in theoretical as well as in practical work. Inspection of the records of those students rating below 110 failed to show that any such students were unable to meet the academic standards of the training school.

*2. Social intelligence as measured by the Moss Social Intelligence Test.*

Faber (2) suggests that social intelligence is essential to success in nursing. In this study our only criteria of success are grades in the theory and grades in the practice of nursing. With this in mind the theory and practice grades for those ranking in the upper and lower quartiles on the Moss Social Intelligence Test were averaged. The average grade in theory for the upper quartile was found to be 86.3 while that for the lower quartile was found to be 82.8. In practical grades the upper quartile averaged 80.6 while the lower quartile averaged 81.2. Note that the upper quartile is only 3.5 points above the lower quartile in theory and is .6 lower in practice. A system of grading using letters instead of figures would probably entirely obscure these differences. We may conclude that social intelligence, using the Moss Test as a measure, is not significant when success in nursing is measured by grades received in theoretical and practical work.

*3. Nursing aptitude as measured by the Moss Nursing Aptitude Test.*

The Moss Nursing Aptitude Test was the only test administered to both the first-year students, all of whom are high school graduates ranking in the upper half of their respective classes, and to the third-year students, only half of whom are high school graduates.

This test was administered to the third-year students since it was desired to correlate Moss Aptitude scores

with theory and practice grades taken over a longer period of time than was possible for the first-year group. This greater accuracy was particularly desirable for the Moss Test because it professes to be designed specifically for nurses for the purpose of predicting an "aptitude" for, and hence success in, nursing.

The first-year and third-year students were compared on the basis of this test. Table II shows the results.

TABLE II

*Comparison of First-Year Group with Third-Year Group*

| Groups | Means | Standard Deviations | Coefficients of Variation |
|---|---|---|---|
| First-year | 116.7±3.2 | 31.16 | 26.7 |
| Third-year | 105.3±3.2 | 31.49 | 29.9 |

The difference between the means of the two groups is 11.4. The critical ratio is 2.5. This means that there are 95 chances in 100 that the difference between the means is a true one.

The writer was of the opinion that the Moss Aptitude Test measured intelligence rather than nursing aptitude. For that reason the partial correlation technique was applied. It was found that when intelligence was held constant the correlation between Moss scores and theory grades dropped from .62±.06 to .26±.0938. The correlation between Moss scores and practice grades dropped from .065±.15 to –.295±.0915. The small degree of relationship indicated between theory grades and Moss Test scores is vitiated by the large probable error. The Moss Aptitude Test seems to be a better measure of intelligence than aptitude for nursing. If this is true, it means that the higher academic entrance requirements have operated to select more intelligent students, since the first-year group scored 11.4 points higher on this test than the third-year group—a group admitted to the training school under more lax academic requirements. Further,

since the Moss Test is so far the only test designed to prognosticate success in nursing, the results yielded by the application of the partial correlation technique indicate that a true nursing aptitude test has not yet been constructed.

*4. Personality as measured by the Bernreuter Personality Inventory.*

The ideal personality test for nurses would measure a number of traits such as attention to detail, conscientiousness, respect for authority, attitude toward criticism, professional attitude, reliability, tact, and other traits deemed essential to success in practical nursing. Such a test for nurses does not exist. This means that for the present at least we are not able to work towards solving the problem of selection of students by administering personality tests to applicants. We can, however, make use of a test such as the Bernreuter in guiding the student nurse into that line of nursing for which she is best fitted. A student ranking high on the Bernreuter scale which measures neurotic tendencies should probably not be advised to follow psychiatric nursing. A student scoring unusually low on the dominance-submission scale of this test would probably be more successful at bed-side nursing than in an executive position. Case studies which were made of those students making extreme scores seemed to indicate that the Bernreuter Inventory might well be used for the purposes suggested here.

The value of the Bernreuter scores for predicting grades in practical nursing was studied by comparing the average grades in practical work of the upper and lower quartiles. The difference between the averages of practice grades for the upper and lower quartiles was found to be .9 in neurotic tendencies and 1.3 in dominance-submission. The higher average in each case was made by the upper quartile, but the differences in

the averages are so slight that a letter system of grading would obscure them.

*5. Correlations obtained.*

The coefficients of correlation obtained between the test scores and grades in theory and in practice may be seen in Table III.

TABLE III

*Coefficients of Correlation Obtained in This Study*

| Tests | Correlation with Practice Grades | Correlation with Theory Grades |
|---|---|---|
| Binet-Simon | −.06±.1 | .368±.08 |
| A.C.E. Intelligence Test | .081±.096 | .618±.061 |
| Moss Social Intelligence Test | .004±.1 | .30±.09 |
| Moss Aptitude (First-year Group) | .065±.1 | .62±.06 |
| Moss Aptitude (Third-year Group) | .061±.15 | .61±.74 |
| Moss Aptitude (Intelligence Constant) | −.295±.08 | .26±.09 |

None of the tests used in this battery proved to be of predictive value so far as grades in practice are concerned. The A.C.E. and the Moss Aptitude seem to be of equal prognostic value for theory grades, according to the table above. However, when A.C.E. scores were re-correlated, dropping the scores of four students not present for the Moss Aptitude Test, the resulting coefficient of correlation was .74±.045. This indicates that the A.C.E. Test is superior to the Moss Aptitude Test in predicting grades in theoretical work.

*6. Comparison of student nurses with college freshmen.*

The first-year group was compared with college freshmen on the basis of the A.C.E. Group Intelligence Test, the Moss Social Intelligence Test, and the Bernreuter Personality Inventory. (Tables IV, V, VI.)

TABLE IV

*Comparison of Student Nurses with Freshmen of University X on the A.C.E. Intelligence Test*

| Groups | Means | Standard Deviations | Coefficients of Variation |
|---|---|---|---|
| Freshmen | 143.9±5.3 | 55.05 | 38.26 |
| Nurses | 109.9±4.2 | 43.35 | 39.45 |

Table IV shows the first-year group as compared with the freshmen of University X, an urban university. Inspection of the table shows that the variability relative to the mean is practically the same in the two groups. Note also that the mean for the freshmen is 34 points higher than that for the nurses. This difference was found to be significant, the critical ratio $\frac{D}{(\sigma \text{ diff.})}$ being 5.04. We may conclude that the freshmen class of this particular university is superior to the first-year student nurses of the training school in which this investigation was made when scores on the A.C.E. Test are the basis of comparison.

TABLE V

*Comparison of Scores Made by Student Nurses on the Moss Social Intelligence Test with Norms for College Freshmen*

| Centiles | Scores on Moss Test | |
|---|---|---|
| | Nurses | Freshmen |
| 75 | 103.75 | 116.5 |
| 50 | 89.50 | 102.0 |
| 25 | 74.05 | 89.5 |

Table V shows that the first-year student nurses scored consistently below the norms for college freshmen on the Moss Social Intelligence Test. Note that the 25 centile is 15.45 less for the nurses, the 50 centile is 12.5 less, and

the 75 centile is 12.75 less than the norms for college freshmen.

TABLE VI

*Comparison of Scores of Student Nurses with Norms for College Women on the Bernreuter Personality Inventory*

| Centiles[1] | Bernreuter Test Scores | | | | | |
|---|---|---|---|---|---|---|
| | Neurotic Tendencies | | Self-Sufficiency | | Dominance-submission | |
| | Nurses | College Women | Nurses | College Women | Nurses | College Women |
| 75 | 62 | 17 | 43 | 47 | 47 | 82 |
| 50 | 11 | –40 | – 2.5 | 8 | 2 | 34 |
| 25 | –35 | –100 | –25 | –30 | –30 | –10 |

In Table VI note that student nurses tend to score higher than college women in neurotic tendencies. The college women, in general, score higher in self-sufficiency and in dominance-submission. The lower scores of the student nurses in dominance-submission may probably be partially explained by the fact that those entering the training school were such persons as were willing to conform to the rigid, almost militaristic discipline of the hospital.

## V. Conclusions

1. On the basis of the Stanford Revision of the Binet-Simon Test, the first-year group of student nurses is for the most part of normal or below normal intelligence, 55 per cent being "normal" and 32 per cent "below normal." Further, at the time of this writing the grades of the "below normals" did not indicate in any instance that these students were doing failing work.

[1] The scores for these centiles were interpolated from the norms for college women. Although these measures are not exact nor absolute, they are probably sufficiently accurate for the cursory comparison made here.

2. In this study the IQ but little influences success in nursing when success is measured by grades in theoretical and practical work. The students with "normal" IQ's rate .6 higher in theory and 1.7 less in practice than the "below normals." Under present conditions in nursing education, a high IQ does not seem to be essential to success. The re-examination procedure used in the training school and in the State Board Examinations probably makes it possible for low grade individuals to become registered nurses. Under the present lax requirements in nursing education, it is unnecessary to bar applicants with "just normal" or "below normal" IQ's from the training school with the excuse that such persons cannot complete the course successfully.

3. University freshmen are superior to the first-year group of the training school when scores on the A. C. E. Intelligence Test are used as the basis of comparison. The mean score of the freshmen studied is 34 points higher than that of the training school group.

4. Of the battery of tests given in this study, the A. C. E. is of greatest predictive value for grades in theory, yielding a correlation of .74±.045. None of the tests proved to be of predictive value for practical work. The coefficients of correlation obtained between intelligence test scores and grades in practical work would seem to indicate that there is practically no relationship between intelligence and grades in practical work. These results are not peculiar since the items on which the student is graded on the Efficiency Record are such that personality traits rather than intelligence *per se* determine the grade in practical work.

5. Social intelligence, using the Moss Social Intelligence Test as a measure, was not found to be significant in predicting success in nursing when success is measured by grades in theoretical and practical work.

6. The Moss Aptitude Test is of fairly high predictive value for grades in theory but in this study was valueless in predicting grades in practical work. This test seems to measure intelligence more than it does aptitude for nursing, the correlation coefficients between the test scores and theory grades dropping from $.62 \pm .06$ to $.26 \pm .09$ when intelligence was ruled out by a partial correlation technique. Similarly the correlation between the test scores and practice grades dropped from $.065 \pm .15$ to $-.295 \pm .0915$.

7. On the basis of the Moss Aptitude Test, the first-year group which is comprised of high school graduates from the upper half of the class is superior to the third-year group, the members of which have had from 1 to 4 years of high school work. This suggests that the higher entrance requirement seems to have operated to select students superior to those admitted under more lax requirements.

8. The scale of the Bernreuter Test which measures neurotic tendencies may be used to help determine fitness for psychiatric nursing. The scale measuring dominance-submission may be used to help determine fitness for executive positions.

9. The student nurses in this study rank higher in neurotic tendencies than college women. The college women rank higher in self-sufficiency and in dominance-submission.

10. The two-fold division of the training necessitates the use of tests which will evaluate the applicant on those traits essential to success in *practical* work as well as in *theoretical* work. Any reliable intelligence test such as the A. C. E. may well be used for predictive purposes in the theoretical aspect of the training course. A test measuring accurately attention to detail, conscientiousness, respect for authority, attitude toward criticism,

professional attitude, reliability, tact, and all those major traits essential to success in practical work would be invaluable in helping solve the problem of selection of students.

## BIBLIOGRAPHY

1. Earle, Mary Goodyear. "Intelligence Testing of Probationers." *American Journal of Nursing.* 1923. Vol. 23, pp. 866-869.

2. Faber, Marion S. "Mental Tests and Measurements." *American Journal of Nursing.* 1928. Vol. 28, pp. 265-271.

3. Hyman, Anna and Dreyfuss, Ruth. "How Intelligent Should Our Nurses Be?" *American Journal of Nursing.* 1930. Vol. 30, pp. 490-494.

4. Metcalfe, Louise. "Achievements of Nurses in Relation to Intelligence Ratings." *Proceedings, National League of Nursing Education.* 1928. pp. 174-175.

5. McPhail, Andrew H. "Psychological Tests for Student Nurses at the Rhode Island Hospital." *American Journal of Nursing.* 1929. Vol. 29, pp. 203-206.

6. South, E. B. and Clark, G. Y. "Some Uses of Psychological Tests in Schools of Nursing." *American Journal of Nursing.* 1929. Vol. 29, pp. 1495-1498.

7. Terman, Lewis M. *The Measurement of Intelligence.* Houghton Mifflin Co. 1916. pp. xviii-362.

8. Young, H. H. "Intelligence Ratings and Success of Nurses in Training." *Journal of Applied Psychology.* 1924. Vol. 8, pp. 377-390.

# THE RELATIONSHIP BETWEEN PLAY AND INTELLIGENCE

PAUL L. BOYNTON
*George Peabody College for Teachers*
and
F. A. FORD
*State Teachers College, Nachitoches, Louisiana*

This study was worked out with fifth grade children located in several different city and parish schools in Louisiana. The intelligence quotients are averages derived from the Otis Self-Administering (Intermediate) and Detroit Alpha tests. Over 1,000 children were tested in a more general analysis,[1] and from this large group were selected the 52 cases herein considered. Each child who appears in the high intelligence group, that is, has an IQ that is above $Q_3$ for his own class, is matched by a child in the low intelligence group, that is, a child with an IQ below $Q_1$ for his class group, who was in the same school and with the same teacher.

As regards the IQ selections it will be seen that a minimum of two quartile deviations separated each pair of children, but that these quartile deviations were determined for each particular grade group studied and not for the more than 1,000 children as a whole. As a result it should be borne in mind that the determination of "high" and "low" were for each particular group of children and not for the group in its entirety. Thus a child who was rated as "high" in one school group would not necessarily achieve the same rating in another school group. Highness and lowness then are relative terms as here employed. While this technique may make for

[1] Ford, F. A., *The Ratio of Achievement to Ability as Found Among Fifth Grade Pupils.* Contributions to Education No. 94, George Peabody College for Teachers, Nashville, Tenn., 1931.

a certain amount of confusion in our first thinking on the problem, it has the advantage of holding fairly constant, by empirical means, certain influences of specific scholastic experiences on the test results. It is almost indisputable that various academic experiences influence group test scores. It was hoped that these influences would be reduced to a minimum by this method of individual selection. The median IQ for the high group is 120.5 and for the low group 83.5. The IQ range for the high group was from 109 to 128 and for the low group from 64 to 95. These data indicate quite definite group dissimilarities, at least.

The following instructions were given to the teachers who agreed to cooperate in the study:

In connection with the special study now being conducted, it is desired to get the following record from every child in the fifth grade who took the five tests recently. Please observe the following suggestions in getting the information.

1. Select a normal week beginning about April 26 for the record.

2. On Friday tell the children that the record will begin on the following Sunday and give each a Pupil's Time Record Card. Ask each child to keep on his card a record of the time he spends in the six different ways indicated on the card.

3. Insist that he spend his time as usual, making no changes in practice for the sake of the record. Advise that the record is to be accurate and that it will not affect his standing.

4. Beginning Monday and running through the following Monday assist each child daily in getting his record as accurate as possible. Make a duplicate record daily for each child as you assist him and keep the duplicate in your possession.

5. As soon as all records are complete, turn over the teacher's copies to the Supervisor to be sent in. You may collect and keep on file the pupils' records for future reference.

6. *Mental Recreation.* Under this point record the time devoted to reading for pleasure; attending picture shows; playing checkers, rook and all similar games; playing the piano or other instruments for pleasure and entertainment; and listening to the radio.

7. *Physical Play.* Include here all physical activities intended primarily for pleasure such as playing ball, kite, tennis, dolls, etc. Fishing, picnicking and similar activities should be included. (Instructions beyond this point not pertinent to this article.)

After the records were sent in it was necessary to eliminate a large number of children whose records would have been considered but for manifest failures to abide by the regulations. The 52 children herein considered had records which were filled carefully and completely in so far as could be determined. In the following discussion of these records all statistical measures have been computed on the assumption that the individual scores form parts of continuous series.

As we introduce our analysis of the records our first consideration may well be centered on the total time spent in play during the week in question. The high intelligence group spent a mean time of 1,873.08 minutes in play as compared with a mean time of 1,538.46 minutes spent by the low intelligence group. In terms of hours, this is about 31.22 hours compared with 25.64 hours; or the high group was spending a little more than 5½ hours a week more in play, on the average, than was the low intelligence group. In still other terms this means that the average pupil in the high intelligence group was spending between 45 and 50 minutes a day more in play of some form, as here defined, than was the pupil of low intelligence. Median averages, 1,800 minutes for the high group and 1,450 minutes for the low group, point to practically the same relationships between the two groups as regards the relative times spent in play.

It is impossible to demonstrate unequivocal mathematical reliability for these differences due to two factors. In both groups the cases included in this analysis are comparatively few in number, 26 for each group, as previously indicated. This tends to make all mathematical errors large and resulting reliabilities low.

Again, individual differences within the groups are very pronounced and the measures of dispersion are large. Thus, the standard deviation of the distribution of the high group's play scores is 805.90, and of the low group's play scores 585.29. The corresponding quartile deviations are 645.84 and 475.00. Incidentally, it should be noticed that individual differences in the amount of time spent in play are more pronounced, or intra-group heterogeneity is greater, among the cases composing the high intelligence group than among those of the low intelligence group, judged by either the standard or quartile deviation methods of measurement.

In furtherance of the idea that any tendency which we have noted for those of high intelligence to spend more time in play than those of low intelligence is purely a group tendency and not one with individual or clinical significance, we should note the following. Among the high intelligence group's children 34.6 per cent of them spent less time in play than did the median child of the low intelligence group, and, in fact, 15.4 per cent of the high group's students spent as little time in play as did those in the lowest play fourth of the low intelligence group. Conversely, 34.6 per cent of the low intelligence group's students spent more time in play than was spent by the average (median) high intelligence case, and 7.7 per cent of the low group's students spent more time in play than 75 per cent of the students in the high intelligence group. While these various per cents of overlapping most obviously diverge from normative expectancies and thus confirm the idea of inter-group dissimilarity, to a certain degree at least, they also confirm the point made earlier in this paragraph that individual diagnosis on the basis of the knowledge of the score placement on one variable—*i. e.*, predicting a child's position in one of these variables from a knowledge of his position in the other variable—would be so precarious as to be wholly unjustifiable.

Next, let us analyse the play of the two intelligence groups from the point of view of type or kind. Here we will refer to mental recreation and physical play in the terms used in the sheet of instructions which was sent to the teachers. As regards physical play we cannot say that there is any particular difference between the two intelligence groups. If the mean is used it would seem that those of low intelligence play physically more, on the average (1,003.85 minutes), than those of high intelligence (976.93 minutes). This condition, however, is reversed rather decidedly when the medians are used. Here the high intelligence group would appear to spend 1,000 minutes a week in this type of play, on the average, as compared with 900 minutes comparably spent by the low intelligence group. Without doubt this ambiguity in averages results from the nature of the physical play time distributions of the two intelligence groups. The scatter or heterogeneity of these scores of the low intelligence group is much more pronounced than with the high intelligence group. The sigma in the first instance is 516.60 and in the second instance 405.35, while the corresponding Q's are 370.00 and 262.50. Still further, we can note that this divergence in scatter tendencies is due to conditions in the upper halves of the groups, in the main. Thus $Q_1$ for the high intelligence group is 650.00 and for the low intelligence group 610.00; the corresponding medians are, once more, 1,000.00 and 900.00; the $Q_3$ points, however, are 1,175.00 and 1,350.00, and thus indicate greater group dissimilarity than is seen elsewhere. This may mean, and it is only one possible suggestion, that among those of lower intelligence if a child spends more than about the average amount of time in physical play he has a tendency to carry it to excess. This should not be taken as a generalization but as a suggestion for further analysis. In any event our data as regards time spent in physical play are not sufficiently pointed to enable us to make a satisfactory or really

justifiable differentiation between the two intelligence groups.

Let us turn now to the last phase of the study, that concerning mental recreation according to the previously stated definition. Here for the first time do we find really appreciable differences between the two intelligence groups. The high intelligence group spent an average (mean) of 919.20 minutes in mental recreation during the week whereas the low intelligence group spent 503.85 minutes on the average in these pursuits. This would indicate that the average child of high intelligence was spending about an hour a day more in mental recreation than was the average child of low intelligence. It should be noted that these mean averages, despite the factors of individual differences and relatively few cases, are so dissimilar as to make for a mathematical reliability of the difference. There is, as can be seen, a mean difference of 415.35; the standard error of this difference is 137.93; the resulting index of reliability is 3.01, and thus the commonly accepted standard, that to be mathematically significant a difference must be at least 3 times its standard error, has been fulfilled. The median difference is both less and relatively less than the mean difference, but it is still quite large and fairly clear in its meaning. It is 370 with a probable error of 105.38. This gives an index of reliability of 3.51 whereas 4.00 is the general criterion before there is acceptance when one is using medians and probable errors. Before discarding this median difference because of its smallness as measured by its error, attention should be given to the fact that the chances[1] are probably slightly more than 99 out of 100 that a true difference does exist and in the direction here indicated.

Though the differences which have just been presented point to a very definite disparity between the groups and

[1] Pearson, Karl. *Tables for Statisticians and Biometricians;* Cambridge University Press; 1914.

though it would appear that one has ample reason to believe that when working with comparably selected and measured groups he will find that the group of high intelligence spends more time in mental recreation than the group of low intelligence, it should be remembered that this is a tendency for the groups only and that it is not sufficiently pronounced to justify individual prediction or diagnosis. Thus 15.4 per cent of the high group is in the lowest 25 per cent of the low group as regards time spent in mental recreation, and 23.1 per cent of the high group's scores fall within the lowest 50 per cent of the scores made by the low group. Conversely, 3.8 per cent of the scores (1 case) made by the low group lie within the highest fourth of the scores made by the high group, and 15.4 per cent of the children in the low intelligence group actually spent more time in mental recreation than did the average (median) child of high intelligence. It is interesting to note that of the four cases who constituted this 15.4 per cent two had IQ's of 69 and 64, and as a result were the lowest and third lowest cases involved in the study. Still further while the children who spent the smallest amounts of time in mental recreation, 0 and 60 minutes, had IQ's of 77 and 79, respectively, the child who spent the next smallest amount of time in mental recreation, 160 minutes, had an IQ of 123.

One other approach may be taken to this problem. When two variables are consistently related the coefficient of correlation should be raised by the elimination of the middle scores of either group. This would result in a correlation of extremes and should accentuate the relationship so appreciably as to make for a spurious coefficient, in actuality. When the high and low IQ cases in this study were thrown into one group and the correlation was run between IQ and amount of time spent in mental recreation the coefficient was found to be $.365 \pm .08$. Hence, it is evident that the relationship is

low for the group thus combined, and when we compute the coefficient of alienation from this coefficient of correlation and find it to be .931 we see how truly precarious would be an individual prediction from one to the other variable.

## Summary and Conclusions

1. This study deals with the play activities of children who are separated from each other in IQ's by at least two quartile deviations of their own groups. Further, it is an analysis of the play activities of relatively bright and relatively dull children.

2. As regards the total amount of time spent in play, the average bright child was found to spend about 45 to 50 minutes a day more in play than does the average dull child.

3. There was no appreciable difference in the average amount of time the two groups spent in physical play.

4. The average bright child spent more time in mental recreation than did the average dull child. This difference amounts to approximately an hour a day.

5. All relationship tendencies which have been noted refer to the groups only and in no manner justify an assumption that the knowledge of an individual's intelligence will enable one to predict either the type or amount of play of the individual, or the reverse. The tendencies are purely group tendencies and though they may have general social and educational import they possess no clinical or individual predictive value.

# AN INVENTORY OF ENGINEERING MOTIVES

H. S. ROGERS AND G. W. HOLCOMB
*Oregon State Agricultural College*

The effective counseling of incoming students and the effective differentiation of their aptitudes and interests in engineering have received considerable attention from engineering faculties in recent years. Solutions for the problem of guiding college students have been suggested or on trial, but to date no one method has proven entirely satisfactory. The study of human nature and of people remains a complex problem. The total integration of particular conditions which exists around the individual at any time includes a great number of variables which are common to all individuals, such as the quality of intelligence, emotions, health, moral character, aptitude, personality traits, and the large and relatively complex group of motives.

That our engineering colleges enroll many who are of mediocre quality or who would "prefer to do something else," is a condition which is unwholesome both from the student's and college's standpoints. As a result of enrolling such students the engineering schools and also other schools have a high mortality between the freshman and senior years. This is a condition which all would like to correct. The various methods by which the problem has been attacked are well known. The generally accepted, by limited methods, are intelligence tests, aptitude tests, and vocational interest tests, each type of which serves its purpose with some degree of effectiveness.

The purpose of the study described herein was to develop an instrument by which a counselor may recognize the dominant motives which describe the attitudes and actions of engineering students prior to enrollment in engineering and which will indicate in a comparative

way the intensity of each motive. The traits and scales for engineering motives presented herein were outlined by H. S. Rogers of the School of Engineering at Oregon State College, and the development and tests were made by G. W. Holcomb.

Intellectual motives are of internal, natural origin and are subjective. Those related to college work in engineering may all be instinctive tendencies of childhood, modified and varied through school experiences. They are colorings or manifold aspects of the one driving force, life. H. E. Bennett in *School Efficiency* says, "The child that is 'full of life' is full of motives and full of activity," and also ". . . any attitude, habit, interest, or other acquired tendency which is effective for directing or arousing conduct of any sort may be regarded as an impulse or motive." Previously expressed motives should be recognized in college, and students with interests and tendencies opposed to engineering should not be enrolled in engineering. Strong motives opposed to engineering training will eventually drive such away from school.

The engineering motives which drive an individual to do certain elementary things with thorough enjoyment should be the dominant ones of a boy anticipating engineering training. The expert cabinet maker finds delight in constructing a fine piece of furniture. The mathematician will spend hours with an abstract equation and find true pleasure when the correct solution is discovered. Yet, in general, the mathematician will receive no thrill in an attempt to build a cabinet nor will the cabinet maker receive a thrill from a "useless" mathematical equation. Likewise, all of the major divisions of academic study, if analyzed, may show some dominant motives which produce the industry required for the successful pursuit of each activity.

Both of the individuals mentioned above have independent and complex motives or incentives for their

work. One may have a motive of craftsmanship, and the other, a motive pertaining to abstract ideas. These same motives were probably in evidence during childhood although in more elemental form. The skilled cabinet maker when a child probably enjoyed working with tools. He was probably satisfied to work alone and found pleasure in constructing articles of use and beauty. The mathematician, on the other hand, enjoyed the problems in arithmetic and algebra. The most difficult problems were probably always a challenge to him and he may have taken pleasure in challenging other students with the "nut-crackers."

A study of engineering students indicates that those who are successful in engineering education have some dominant motives which students in commerce, for example, do not have. It would not be expected that a student of commerce would be intensely interested in problems of scientific research; he may, however, have a strong motivation toward transactions of barter and trade. Thus, in all practical fields of engineering, agriculture, medicine, commerce, home economics, etc., there is undoubtedly a small group of distinctive motives predominating each and comparatively independent of the other fields. These dominant motives afford the drives which the student manifests in his own personality. If the group of motives which predominates in one student happens to find opportunity for expression in engineering, the student should select engineering; likewise, the student with commercial motives should find enjoyment in business training.

## Engineering Motives

The major interests and instincts which may find expression in engineering training and practice and which will motivate students and practitioners alike seem to be included in the following groups:

1. Constructive instinct.
2. Scientific curiosity in mechanisms.
3. Abstract ideas of a quantitative nature.
4. Precise observation and description of physical things.
5. Practical application.
6. Planning and organization.
7. Manual activity and craftsmanship.

*The Constructive Instinct.* The advantage that definite constructive ability contributes toward accomplishment in engineering work can hardly be questioned. The continual play with the tools, the putting together and taking apart of things gives a boy an accumulated store of knowledge of the material world.

William James in *Talks to Teachers* says, "Constructiveness is instinct most active." For engineering we desire students in which this instinct is highly developed. We should search for students, who in pre-college days were continually demonstrating a constructive instinct, which James says, "accumulates a store of physical conceptions which are the basis of his (the student's) knowledge of the material world through life." The boy who frequents the blacksmith shop, garage, and the carpenter's shop, and enjoys building wagons and mechanical devices, acquires a tendency to build and construct things which carries over into manhood.

*Scientific Curiosity in Mechanisms.* Curiosity gives us innumerable expressions, but one kind of scientific curiosity finds expression in a general field which is closely related to engineering. A boy observes a piece of mechanism in a garage and wonders why, or what, or how, and reaches out as intelligently as possible with the partial knowledge which he may possess to satisfy his curiosity. His inquiring mind is always aroused by

mechanisms of a scientific nature which he does not understand. Such is scientific curiosity in mechanisms, and the pre-college boy who is aroused by a curiosity of this nature is a prospective student of engineering science.

*Abstract Ideas of a Quantitative Nature.* The term, abstract ideas, may be explained as one phase or aspect of an object based upon concepts distinct from the objective features. Good engineering students seem to have an incentive for work with abstract ideas of this kind. Even eminent engineers enjoy challenging their associates with "brain stretchers." The boy who is motivated by a desire for such abstractions will, likewise, enjoy challenging his fellow students with problems which are difficult or new. Any boy who has a tendency to work with things which are obscure and removed from ordinary physical contacts, yet which have the very foundation of their transcendental nature in the material world, such as the difficult problems of a mathematical and geometrical nature encountered in the high school, has a motive favorable to engineering.

*Precise Observation and Description of Physical Things.* The inclination to make accurate quantitative measurements of facts and to carefully describe phenomena, by which changes and relationships are observed, is necessary in engineering work. Likewise, work associated with the physical sciences requires that measurements be taken which are accurate within certain prescribed limits of error. Opinion and vague terms, such as *more or less, much and little,* when gauging likenesses and differences, are not adequate for measures and descriptions when designing structures, writing specifications, or working in the research laboratory. The engineer must know just how much better or worse one piece of steel is than another. He must be precise in all of his physical observations and descrip-

tions. The high school boy who spends his leisure time reading and comparing the specifications of automobiles or other machines with a critical attitude, or who goes to a machine shop to investigate calipers and other accurate measuring devices, has a motive of precise observation and description which is a characteristic of practicing engineers.

*Practical Application.* The purpose of a considerable portion of engineering work is the practical application of theoretical ideas and mechanisms. If a new piece of research work is announced, professional engineers interested in the field immediately inquire for its practical and commercial use. Some will perform experiments of their own to verify the statements made regarding the applications of new theory. Likewise, the engineering freshman, who as a high school student was always testing and putting new ideas into service around the home and school, has this qualification for engineering work.

*Planning and Organization.* The value of an interest in planning and organization cannot be over-estimated as a valuable asset to engineers in all divisions of their work. The research man plans his laboratory procedure, the designer lays out and organizes his methods of computation, and the construction engineer plans and organizes the work on the site of the project, in fact, the prophets and fathers of scientific management were engineers. This ability to plan and organize work is manifested in early school days by the boy who keeps his tools in order on the work bench and who likes in advance to plan and organize the procedure of his work.

*Manual Activity and Craftsmanship.* The Vocational Interest Blank prepared by Edward K. Strong, Jr., shows by the weights on the engineer's scale that engineers as a whole have considerable interest in the manual arts. Men who recognize perfection in the operation of

both hand tools and machines generally feel a true pleasure when operating them themselves. This pleasure found with the use of tools and machines is manifested in pre-college days by the boy who frequents the manual arts shop of the high school on Saturday afternoons or on holidays in order that he might work with the tools. Boys who are successful in engineering classes are generally interested in making devices which require some degree of skill and craftsmanship.

The foregoing analysis of dominant engineering motives provides a foundation upon which an inventory of incentives and motives was developed for determining the extent to which successful engineering students gave expression to these prior to their college matriculation.

## The Inventory Blank

The "Inventory of Engineering Motives" was developed as illustrated by the sample form. Each motive was divided into five qualifying graduations of intensity of satisfaction or enjoyment which a boy might feel when giving expression to the motivation. The motives are not named in the inventory but instead statements qualifying them are made in order that a person marking the blank will not necessarily know the purpose or value of any particular statement.

Each group of qualifications for a motive is arranged, however, with either increasing or decreasing intensity of pleasure. For example, the first five statements of the blank are designed to measure the value of the incentive of constructive instinct in an individual. The statements carry the reader from a feeling of repulsion toward the handling of machines, to a feeling of thorough enjoyment when working with them. Likewise, the statements number six to ten inclusive classify scientific curiosity. In a similar manner the other motives are graduated by groups of five throughout the blank. The

inventory may be made with few or no oral directions as desired. The directions at the top are sufficient for the students' information. It seems advantageous, however, to read the directions orally to students in order that the counselor may be positive that all understand them. No limiting time is required for finishing the blank; fifteen to twenty minutes should, however, suffice for the slowest student.

## Scoring

The inventory is scored by means of the form shown in Figure 1, on which the tendencies to motivation are designated by names.

### Inventory of Engineering Motives

Name_____ Class_____ Dept_____

| | (Fr., Soph., etc.) | | | (C. E., E. E.) | |
|---|---|---|---|---|---|
| Constructive instinct | 1 | 2 | 3 | 4 | 5 |
| Scientific curiosity in mechanisms | 10 | 9 | 8 | 7 | 6 |
| Abstract ideas of a quantitative nature | 11 | 12 | 13 | 14 | 15 |
| Precise observation and description of physical things | 20 | 19 | 18 | 17 | 16 |
| Practical application | 21 | 22 | 23 | 24 | 25 |
| Planning and organization | 30 | 29 | 28 | 27 | 26 |
| Manual activity and craftsmanship | 31 | 32 | 33 | 34 | 35 |

Figure 1

The horizontal row of numbers following constructive instinct, 1-5 inclusive, refer to the numbers of the statements on the inventory. The first vertical row of num-

bers, 1, 10, 11, etc., are on the low side of the scale; that is, the side of no motivation, and numbers 5, 6, 15, etc., are on the strong side of the scale or maximum motivation. The middle vertical row may be considered as a modicum of motivation.

A scorer may encircle the numbers on the scoring blank which corresponds to the number checked by the student on the inventory and obtain a profile of the student's motivating tendencies by connecting these numbers with an irregular line. Such a profile furnishes the counselor an accurate survey of the motivating tendencies of a student, which is valuable for use in interviews.

## Characteristics of the Inventory

The justification of the use of the seven motives and the exclusion of all others is predicated upon the hypothesis that they are the important and dominant motives of practicing engineers. Observations of results obtained from the use of the inventory seem to indicate that the selection of the seven motives was valid, although additional motives could be included which are important to particular branches of engineering. The validity of the inventory blank can be established, however, only in terms of its usefulness in vocational guidance and personnel work. Experimental use of the inventory seems to indicate some significant values in diagnosing student interests.

For the purpose of testing the validity of the inventory, it was scored by 204 engineering freshmen during the first two weeks of the college year 1931-32, and for comparison it was scored by 88 commerce freshmen and 25 vocational education seniors and graduate students about the same time. Figure 2 shows the profile of the average scores for each of these groups. The average scores of the engineering group are considerably higher than either of the other groups. The percentage difference in group scores between commerce and engineering

is 22 per cent in the case of "Scientific Curiosity" and 2 per cent for "Planning and Organization." The former is the maximum variation. It is significant to note that lesser differences between the groups occur in two motives, "Planning and Organization," and "Practical Application," than occur in the other five. Students in commerce should in all probability be expected to score as high as engineers in planning and organization which the results verify. The score for vocational education students would be expected to vary from the other two since these students are interested in a variety of subjects, such as engineering, agriculture, mathematics, commerce, and natural sciences.

INVENTORY OF ENGINEERING MOTIVES

Name_____ Class_____ Dept._____
(Fr., Soph., etc.) (C. E., E. E., etc.)

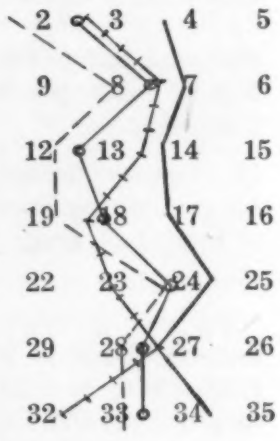

Constructive instinct 1 2 3 4 5
Scientific curiosity in mechanisms 10 9 8 7 6
Abstract ideas of a quantitative nature 11 12 13 14 15
Precise observation and description of physical things 20 19 18 17 16
Practical application 21 22 23 24 25
Planning and organization 30 29 28 27 26
Manual activity and craftsmanship 31 32 33 34 35
204 Engineers
25 Vocational Education
88 Commerce
19 Commerce Women

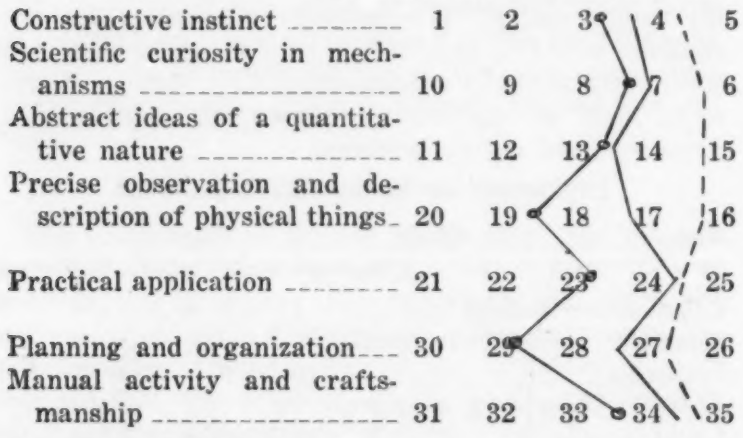
Figure 2

The profile of the average scores for 19 women students in commerce is also drawn in Figure 2. Their

average is below the average of the other groups for the first four motives, the ones which are predominantly scientific and engineering in nature while it is well within the other groups on the last three.

INVENTORY OF ENGINEERING MOTIVES

Name _____ Class _____ Dept. _____
(Fr., Soph., etc.) (C. E., E. E., etc.)


Constructive instinct 1 2 3 4 5
Scientific curiosity in mechanisms 10 9 8 7 6
Abstract ideas of a quantitative nature 11 12 13 14 15
Precise observation and description of physical things 20 19 18 17 16
Practical application 21 22 23 24 25
Planning and organization 30 29 28 27 26
Manual activity and craftsmanship 31 32 33 34 35

19 Engineers transferred o——o
All engineering freshmen ———
33 Engineers, grade average above 1.7 — — —

Figure 3

Figure 3 gives (1) the composite profile of a group of nineteen students who transferred from the School of Engineering to other schools of the institution such as forestry, agriculture, and vocational education, (2) the composite profile of all engineering freshmen, and (3) the profile of 33 engineering students with the highest grade averages.

These 19 students who transferred all received failing grades in one or more engineering courses, which fact was probably the principal reason for the change in their curricula. The profile of this group shows a motivation

which is considerably less than the motivation of the entire group of freshman engineers. The profile for the group of 33 leading engineering freshmen who received grade averages of 1.75 or better for the first term allows a ready comparison of the degrees of motivation which might be expected in good and poor engineering students. Seventy-one per cent of the students who received high grade averages recorded only two motives as average in quality of intensity, and the others above average, while sixty-four per cent, on the other hand, of the group which transferred out of engineering each recorded two or more motives below average in quality of intensity as determined by the middle statement on the rating blank.

It may be reasonably concluded that the inventory blank provides the vocational counselor or personnel officer with a valid instrument by which he can quickly picture a student's engineering motives. When an interview is given, the profile instantly describes the motivating tendencies and saves considerable time for the counselor. It also affords a suitable comparison with other students who are, or are not, successful in engineering study. The inventory may be scored by students at the beginning of the freshman year, before trouble develops, and before any student has "something to hide." After a student finds himself in difficulty he may naturally and unconsciously cover up important motives which might influence a counselor's decision. Continued use of the inventory has shown that a knowledge of engineering motives when used in conjunction with an intelligence test and the high school record of a student is of considerable importance when diagnosing individual cases.

It seems possible that similar inventory blanks could be constructed for other divisions of learning. An inventory for business administration would in all probability include a motive describing in a comparative way the attitude of a student toward problems of barter and trade. The same inventory would, on the other hand, not

include the motives of manual activity and craftsmanship, or scientific curiosity in mechanisms. The limited experience with this inventory of engineering motives suggests wide possibilities for similar tests.

## Inventory of Engineering Motives

Name_____ Class_____ School_____
(Fr., Soph., etc.) (Engr., Com.,etc.)

Directions: The following list of interests and activities describe some of the attitudes and actions of students prior to their enrollment in college. Read them carefully noting the EXTENT of interest or action in all cases and check with a cross those which describe your actions and feelings before your entrance to college:

1. Never *cared* to use tools or handle machinery and avoided them. ( )
2. *Used* tools, *made* home repairs, and *maintained* machinery *ONLY* as it was necessary. ( )
3. Liked to *repair and overhaul* engines and machinery and all things around home. ( )
4. Liked to *build* wagons, mechanical devices, and electrical apparatus and *did build SOME*. ( )
5. Enjoyed *building* machinery, engines, electrical apparatus, and instruments and did so *CONTINUALLY*. ( )

---

6. Took delight in *ALWAYS seeking and finding explanation* of performance of machines, tools, and devices about home, garage, shops, or laboratory. ( )
7. *AT TIMES took* machines and devices *apart* to find out how they worked and sought scientific explanation. ( )
8. *Liked to know* how and why home and auto or farm mechanisms worked. ( )
9. Were *satisfied* with the *explanations* of how and why machines and devices worked as these were *encountered* in duties and tasks. ( )
10. Paid *NO particular* attention to explanations of performance of things and machines and *did not examine* them. ( )

11. Did *NOT* like mathematical explanations and problems. ( )
12. Did assigned mathematical work in average manner *WITHOUT particular* pleasure. ( )
13. *Liked to do* assigned mathematical work but gave *little extra* thought to such things. ( )
14. Were *DETERMINED to work* difficult and tricky problems which were presented by your teachers and associates. ( )
15. Took *delight* in working *hard* problems of mathematical type encountered in school and in finding hard ones to challenge associates. ( )

---

16. *ALWAYS sought and enjoyed* exact and quantitative descriptions of things like specifications of autos, etc., and *enjoyed using* exact measuring instruments. ( )
17. *Liked to get* exact descriptions of performance of airplanes, cars, etc., and *like to work* with micrometers, microscopes, etc. ( )
18. Dissatisfied with general description of things and *wanted measurements* for comparative purposes. ( )
19. Accepted accurate and quantitative descriptions as encountered but *did NOT seek* them out. ( )
20. *Did NOT like* accurate measuring and gave no particular attention to detailed comparison of physical things. ( )

---

21. Believed interests and activities should *NOT be evaluated* in terms of practical use. ( )
22. Gave *NO thought* to *practical* nature of interests in work or play. ( )
23. Had *SOME interest* in practical things but was *NOT particularly satisfied* by their practical nature. ( )
24. Liked to work or play *ONLY* with things that were of *practical use.* ( )
25. Found *MUCH satisfaction* in the *practical* nature of things with which you worked and played. ( )

---

26. Found *much satisfaction* in *organizing* tools and work and in *planning* procedure and *ALWAYS did so.* ( )
27. *Liked to* plan and organize all work and activities and *USUALLY did so.* ( )
28. Outlined and planned *required* work upon initiative. ( )
29. Planned and organized details of work *ONLY* when it was required. ( )
30. *Disliked* planning and organizing work or activities. ( )

31. *Avoided* work requiring manual operations and skill of a crafts nature. ( )
32. Did work requiring manual activity and skill with *indifference.* ( )
33. Worked at *assigned* tasks requiring manual activity and skill *with satisfaction.* ( )
34. *Liked* working with tools requiring skill and action and *did so AT TIMES* for pleasure. ( )
35. Found *MUCH satisfaction* in craftsmanship with hand tools and enjoyed *FREQUENTLY using* machines that required manual operation and skill. ( )

# BROADCASTING VERSUS LECTURING IN PSYCHOLOGY; PRELIMINARY INVESTIGATION

HAROLD V. GASKILL
*Iowa State College*

To determine the relative effectiveness of broadcasting and lecturing to college students, the following experiment was conducted.[1]

One hundred and thirty college students in the experimenter's classes in Psychology served as subjects. These were divided into two groups, the division being made to obtain homogeneity on the basis of (1) scholarship, (2) intelligence, and (3) familiarity with the voice and style of the speaker. Two twenty-one minute talks on Psychology and Athletics were broadcast from WOI at a one week interval. These talks were not especially prepared for this experiment but were on the regular weekly psychology series. For the first talk one-half of the subjects, Group A, listened to the speaker at the studio, and the other half, Group R, listened by means of their own radio receiver. (That is, they used the radio receiver which they ordinarily used. Some of these were in dormitories and fraternity houses, and the number listening to one receiver varied from 1 to 8 in a small sub-group. Those who had no place of "their own" in which to listen were invited by some of the others to use radios available. About 30 per cent were not habitual radio listeners.) For the second talk the groups exchanged position —Group A listened to the speaker by radio, and Group R listened directly to the speaker at the studio. Three days after each of the two talks all of the subjects were given an objective examination covering the subject matter of

[1] Nothing similar in literature, to December 1932. Report was given at Detroit meeting (1932) of Public Speaking Association of experiment involving about 25 persons in which broadcast material was subsequently tested.

the talk. These examinations consisted of true-false statements and completion items. The time interval between the talk and examination was constant for all subjects, viz., three days.

The same examinations were given to three classes of persons who did not hear the talks. Some of the information given might have been common knowledge, the result of coaching, or the result of a general interest in athletics. One class to whom the examinations were given consisted of one hundred twenty-five persons who did not hear the talk. It was thought also that the subject matter might have been hinted at in classes in psychology. Hence, 90 students in psychology, not in the experimenter's classes, were given the examinations. Finally, one group of people who were in the experimenter's classes but who did not take part in the experiment, 38 in number, took the examinations. Thus, if the experimenter had mentioned some of the subject matter of the two talks in class, it would be apparent.

The persons listening at the studio sat on chairs in front of the speaker's table, beyond the microphone. They were instructed not to take notes. Persons listening by radio also took no notes. Both groups were instructed to sit quietly, without talking, during the broadcasts. Those listening by radio were to maintain the same degree of quiet that was observed by the group at the studio. Type of radio receiver used was varied, but 80 per cent of the receivers were equipped with dynamic speakers. It was felt that conditions of reception would be constant, since a local station of 5,000 watts was used.

Scores on the examinations indicate a small but significant superiority of broadcasting. Every individual serving as subject had a higher score on the examination following a radio presentation than on the examination following a studio lecture. In no instance did this vary. The scores for those listening by radio were higher for each broadcast than the scores for those listening directly at the studio. The examinations show an internal re-

liability of .88 (odd-even items). It is quite obvious from a comparison of test scores that the subject matter of the two talks was not common knowledge, was not included in psychology courses, and was not given in the speaker's classroom lectures. Group A, listening at the studio for the first talk, consistently raised its examination scores for the second talk received by radio receiving sets. Group R scores similarly show consistent drop for the second talk (at studio). A complete reversal occurred.

The data of individuals were further analyzed and the following results were found:

1. There is no relationship between frequency of listening to radio programs and excellence in our examination scores.

2. No consistency in favorite radio programs, and favorite types of programs.

3. Thirty per cent of the people had not listened to a radio program since school started in the fall. Less than 20 per cent were daily listeners.

4. To the question, "Where do *you* think you could have 'listened' better, studio or radio?", there was no consistency of preference with excellence. In each group, about half preferred studio.

GROUP A

| | M | $\sigma$ | N | Order |
|---|---|---|---|---|
| Studio | 36.2 | 4.28 | 70 | 1st. |
| Radio | 41.0 | 3.26 | 70 | 2nd. |

D/P.E. diff. = 11.2

GROUP R

| | M | $\sigma$ | N | Order |
|---|---|---|---|---|
| Studio | 34.4 | 3.81 | 60 | 2nd. |
| Radio | 39.5 | 2.86 | 60 | 1st. |

D/P.E. diff. = 8.6

CONTROL GROUPS

| | M | $\sigma$ | N |
|---|---|---|---|
| I. | 26.9 | 4.22 | 125 |
| II. | 24.0 | 3.92 | 90 |
| III. | 25.2 | 3.86 | 38 |

# THE INFLUENCE OF NOISE ON PRODUCTION AND FATIGUE, AS RELATED TO PITCH, SENSATION LEVEL, AND STEADINESS OF THE NOISE

DONALD A. LAIRD*
*Colgate University Psychological Laboratory*

Although it is generally believed that noise cuts into industrial production and increases worker fatigue, there is nevertheless a dearth of quantitative data showing the relationships which exist between the amount or kind of noise and production or fatigue. Where data are on record engineering usefulness is not allowed since the noise conditions are specified in very meaningless although practical terms, such as "the noise was made by pulling tin vegetable cans across a hardwood floor," or "all observers agreed that the noise for the second phase of the experiments had been reduced by at least one-fourth in loudness." (1)

The increasing strength of anti-noise crusades makes it especially desirable to measure the relationship between various factors of noise and production or fatigue with some semblance of engineering precision in the experimental control of the noise. It has been the aim of the present work to make such measurements with the noise factors so controlled and specified that an acoustical engineer can readily interpret a definite industrial situation in terms that are both meaningful and predictable.

## THE INDUSTRIAL SITUATION TESTED

Industrial operations involving work of a dexterous, repetitive nature requiring close eye and hand coordina-

* With the assistance of Allen Beall, Wayne G. Benedict, Jr., Richard Buehl, Morford L. Downes, Ann Getchonis, George Galbraith, Donald K. Hodges, Edwin Ingraham, Donald S. Keller, Arnold Sanborn, and Fletcher Waller.

tion were chosen as most typical and most significant for the tests. An "experimental factory" was built for employing two workers simultaneously. The work period was four and one-half hours, or an entire afternoon. The work was a prototype of a large number, perhaps the largest, of routine light industrial and office operations.

The test workers held an electrical stylus in their right hand and with this stylus touched through small holes as they appeared in a moving tape to establish contact with a metal plate exposed through each hole. The holes, which were irregularly spaced and of irregular size, appeared in a small window directly in front of the worker. Both workers had exactly the same sequence of holes presented by the "miniature factory" simultaneously. There were 10,500 holes to be touched through in the half day's work.

As in a present-day factory the workers had to be in step with the tape which figuratively represents the conveyor bringing parts to be assembled, or turned, or polished, or inspected. Each hole missed or imperfectly touched through represents a piece not inspected, or not assembled, or not operated on according to the maximum opportunity afforded by the system of the plant.

Production was measured by electrical counters. The test room was 10 feet in each dimension, was lighted entirely by artificial light, and was ventilated by an exhaust fan which changed the air every three minutes.

## The Noises Used

The noises used were generated entirely by electricity. Western Electric audiometers were used as the original noise source, and their output was amplified through a special amplifier with impedences arranged to match the output impedence of the audiometers. This made possible amplification without distortion of the character

of the audiometer sounds. A constant electrical check of the energy going into the dynamic loud speaker was maintained by a meter across the speaker input. The noises used are specified later.

## The Workers Tested

Four healthy young men were tested. Two of these worked in the experimental set-up for almost three months in the spring of 1931, the other two worked for the same length of time in the summer of 1931. The data charted are a composite of the records of these four subjects. The workers were paid forty cents an hour from special funds.

## Experimental Precautions

All subjects worked on the dotting machine for the first twelve days under different noise conditions before records of production were used for study; the twelve days were allowed to make certain that the records used for this report were not vitiated by learning or by the subjects "not being used to the noise." Records for humid days, hot days, or days when one of the subjects felt indisposed are also not included in the report.

An incentive for maximum effort was provided at the end of each 500 holes when a two-minute rest pause was taken and the subjects were told which one had the better record for the 500 holes just completed. As a further precaution against the subjects "not being used to the noise" the data for the louder noises were obtained last in the course of the experiment when the subjects should have been adapted to the noise if they were ever going to be used to it.

Two full records were obtained from each subject at each noise level and character used. The points on the charts are thus based upon two records from each of

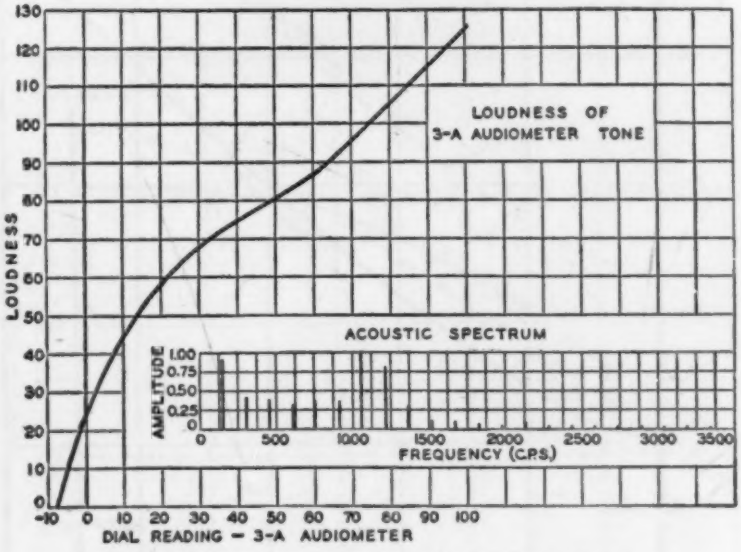
four subjects tested, or a total of eight records for each point.

## Effects of a Steady and Complicated Noise

The buzz type noise of the 3-A audiometer was amplified in the test room for obtaining this data; the frequency characteristics of this noise are shown in Fig. 1.

FIGURE 1

*Frequency and Loudness Characteristics of the Noise of the 3-A Audiometer.*



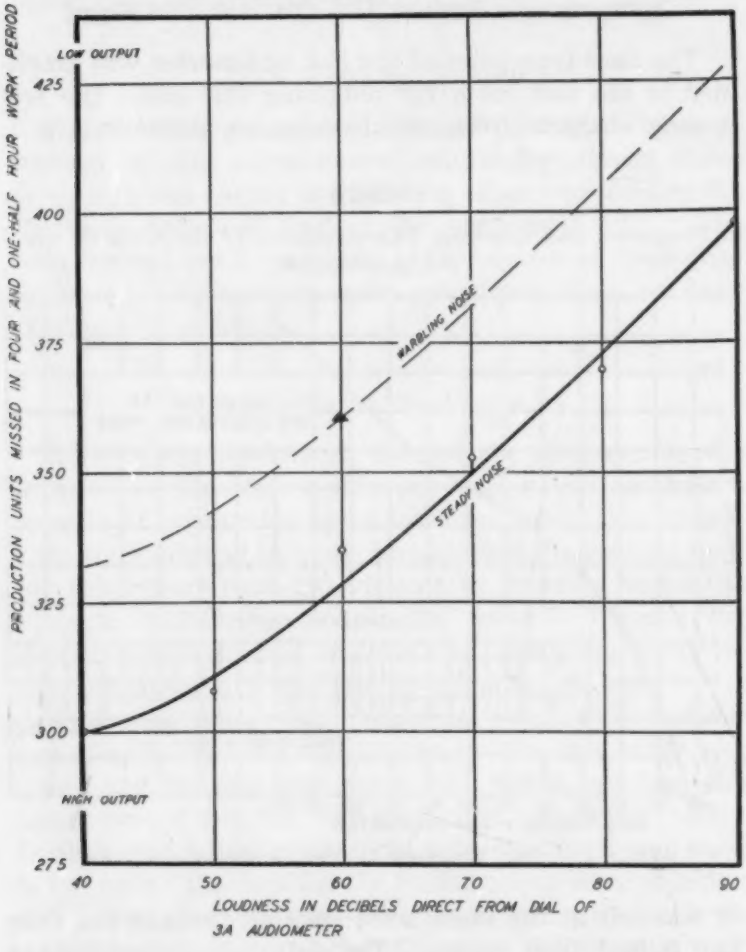

After Fletcher, Courtesy D. van Nostrand Co., Inc.

It was left at the same noise level all through the four and a half hour period. The data for production are shown by the solid line in Fig. 2.

It will be noticed that there is no breaking point beyond which further increases in noise cease to cut into production. In other words, a reduction in noise from

FIGURE 2

*Output at Dexterous Repetitive Work Requiring Close Eye-Hand Coordination as Influenced by the Loudness of a Complex Noise Bathing the Worker*



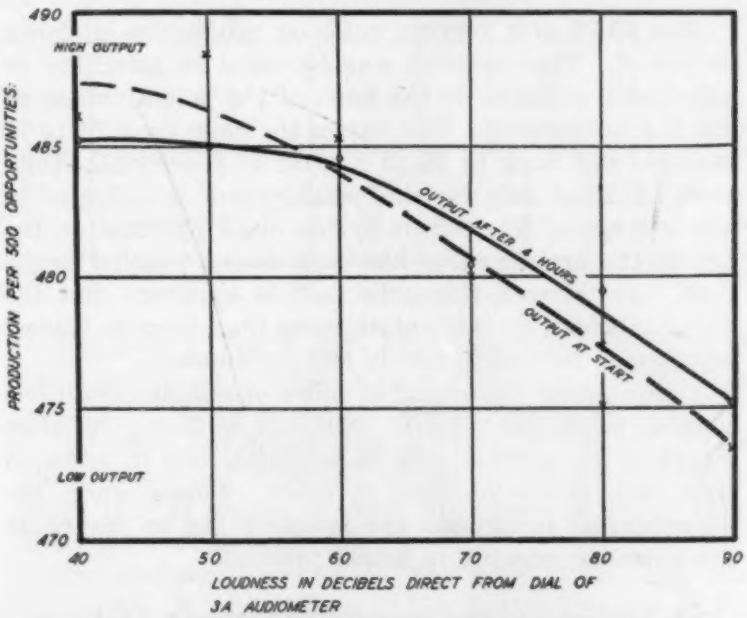

any level tested is characterized by a further increase in output. This gain in production, however, is more marked with reductions from the more intense noise

levels. This is the answer to the belief that some work places are too noisy to benefit appreciably by reducing their noise since they cannot be made acoustically quiet. Merely making a work place quieter is desirable in the light of these data. A drop in loudness of 5 per cent from the loudness level of 70 decibels reduces missed production by 3 per cent.

Fatigue, as shown by changes in productivity during the work period, is shown in Fig. 3. Although these

FIGURE 3

*Fatigue at Dexterous Repetitive Work Requiring Close Eye-Hand Coordination as Influenced by the Loudness of a Complex Noise Bathing the Worker*




changes may be within the experimental errors, at a noise level of 40 and 50 decibels the initial output (for the first 2,000 holes) is greater than the terminal output (for the last 2,000 holes). This is typical of the normal

course of fatigue. At 70, 80 and 90 decibels, in contrast, the reverse holds. This suggests either that the usual warming up phase of work is greatly delayed at these higher intensities, or that the subjects are straining to spend more effort in accomplishing less. The second is confirmed by the introspective reports of the workers and by their sensations of muscular fatigue at the end of the work period.

The fatigue data may be summarized by stating that the records for 40 and 50 decibels show a natural fatigue tendency while the louder levels reveal an abnormal fatigue tendency.

## The Effect of a Varying Noise

The effect of a varying noise on production is shown in Fig. 2. The variation was produced by attaching an adjustable eccentric to the knob of the potentiometer of the 3-A audiometer. This varied the noise from 50 to 70 decibels and back to 50 in a cycle of 8 seconds. Data were obtained only for this point, which is indicated at the average of 60 decibels by the black pyramid on the chart; the broken curve has been drawn parallel to the solid line through this point. It is apparent that the varying noise cuts into output more than even its loudest component does when steady and continuous.

Under actual industrial or office operating conditions varying noises are the rule. Not only do they vary intermittently or aperiodically in loudness, but in addition they also warble or vary in pitch. Consequently the experimental conditions are probably not so severe as are noise components in actual practice.

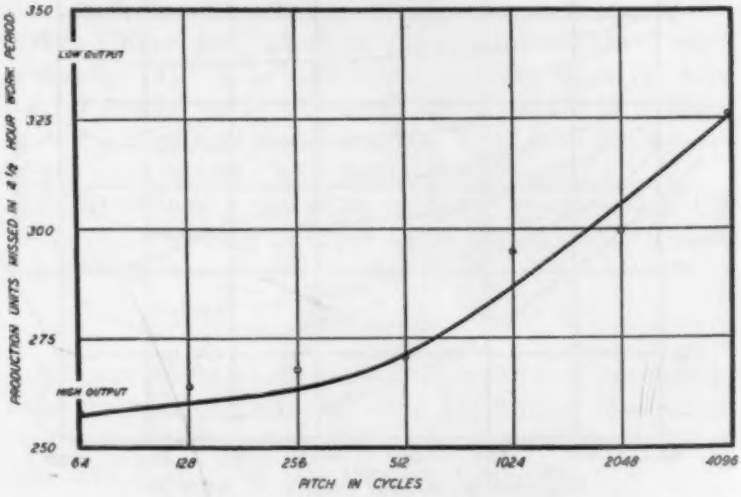
## The Effect of Pure Pitches of Uniform Loudness

The pure pitches were amplified from the 2-A audiometer, held at a uniform loudness level in accordance with the contour lines of equal loudness as specified by

Fletcher (2, page 230). The loudness maintained for all pitches was equivalent to that of 512 cycles at 60 decibels on the audiometer scale.

FIGURE 4

*Output at Dexterous Repetitive Work Requiring Close Eye-Hand Coordination as Influenced by the Pitch of a Pure Tone Bathing the Worker*



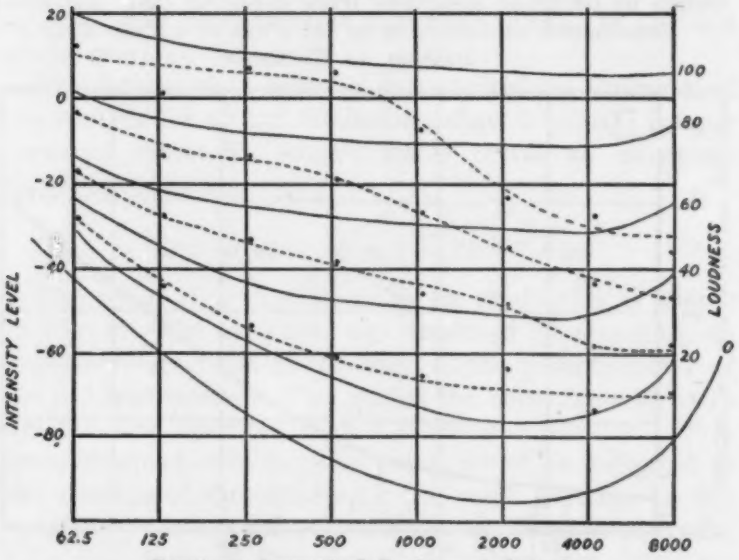

The data in Fig. 4 show that the higher the pitch of the noise the more it cuts into productivity. Squeaks are apparently worse on output than are rumbles.

It is also significant that the data of Fig. 4 parallel those previously reported on the mental experience of annoyance as related to pitch, and shown in Fig. 5 from this earlier experiment (3).

The uniform loudness of the pure tones used was equivalent to the loudness of the 3-A audiometer at the lowest level used in the data of Fig. 2 (2, Fig. 110). Emphasis should be given to the fact that all of the relatively pure pitches reduced output less seriously than the complex noise of equivalent loudness with the exception

of the pitches of 2,048 and 4,096 cycles frequency. The noise of the 3-A audiometer is especially rich in high frequencies, and the reliability of the data is confirmed by the closeness by which the output on 2,048 and 4,096

FIGURE 5




Courtesy Acoustical Society of America

Annoyances as related to pitch. Solid curved lines give contours of equal loudness for pitches indicated at base of chart, after Fletcher (2). Dotted lines are contours of equal annoyance, after Laird and Coye (3). Note that drop in dotted line above pitch 500, representing increased annoyance, parallels rise in curve of Fig. 4, representing decreased output.

cycles matches the output under a corresponding loudness of the 3-A audiometer noise (2, page 232, also see Fig. 1, this report).

## INCIDENTAL OBSERVATIONS

With the more intense noises muscular stiffness was noted, especially in the neck and legs. Neither of these groups of muscles was used during the work period, and

this phenomenon suggests that the general muscular strain accompanying the increased exertion in working under the noisier conditions was accompanied by an accumulation of lactates without the benefits of massage by activity to hasten their dispersion.

Most unexpected and of greatest significance, however, was the great increase in the volume of urine excreted at the level of 80 and 90 decibels. This is in harmony, however, with the work of Corbeille on changes in the volume of the kidney and spleen of dogs under different noise influences (4). This observation is being made the subject of special study and is in line with the hypothesis that the effects of noise are due largely to the appropriate noise causing a biological fear reaction.

At 90 decibels a humming persisted in the ears of the subjects for as long as three hours after the day's work was ended.

## SUMMARY

Measurements of the effect of accurately controlled noise upon production in work duplicating the neuromuscular activities involved by a wide array of industrial and office tasks indicates the following:

1. A varying complex noise has the most serious effect upon production.

2. A complex but steady noise is worse on production than a relatively pure tone.

3. A pitch above 512 cycles cuts into production more than a low pitch of the same loudness; above 512 cycles each increase in frequency is followed by a loss in production.

4. A reduction in a complex noise regardless of the original loudness is accompanied by increased output.

## References Cited

1. Donald A. Laird, "The Effects of Noise," *Journal of the Acoustical Society of America*, January, 1930, Vol. 1, pages 256-262.

2. Harvey Fletcher, *Speech and Hearing*, New York, D. van Nostrand Co., pp. xv, 331, 1929.

3. Donald A. Laird and Kenneth Coye, "Annoyance as Related to Loudness and Pitch," *Journal of the Acoustical Society of America*, 1929, Vol. 1, pages 158-163.

4. C. Corbeille, "Respiratory Responses to Acoustic Stimuli in Intact and Decebreate Animals," *American Journal of Physiology*, 1929, Vol. 88, pages 481-490.

# JAPANESE WORK CURVES

THOMAS R. GARTH, KUNEHEI IKEDA, and
DOROTHY A. GARDNER
*University of Denver*

Some work curves are concave, some are convex, some of the horizontal type and some of the regularly ascending, or of the descending type. It is not certainly known what causes differences in work curves, whether it is sex, skill, differences in age, physical endurance, or race.

The interest of the present study is to ascertain if race makes any difference in the shape of the curve. There have been published already five studies or racial work curves. The first one occurred in this Journal in 1919 and the others have appeared therein from time to time (1, 2, 3, 4, 5). This makes the sixth study with the same materials and the same technique. The races so far studied are Whites, Negroes, Indians, full and mixed bloods, nomadic and sedentary, and Mexicans.

An examination of these studies shows no differences which may be attributed to race. Where differences appear they may be easily accounted for by selection, or nurtural factors. Sedentary Indians differ from nomadic Indians, young Mexicans differ from older Mexicans. The chances are strong for the possibility that Whites differ from Whites according as they are selected.

The purpose of the present study is to inquire into the efficiency of Japanese children, boys, an older and a younger group, as they carried on the continuous performance of a task in addition. The younger group was composed of fifty-two third grade and fifty fourth grade boys, in all one hundred two. The older group was composed of fifty-four seventh grade and fifty eighth grade boys, in all one hundred four. This makes a total group of two hundred six Japanese boys.

The performance of the Japanese children was brought into comparison with groups of White boys of more or less similar composition. The younger group of Whites was composed of fifty-one third grade and one hundred and eleven fourth grade boys, in all one hundred sixty-one children. The older group of Whites was composed of one hundred seventeen in the seventh grade and thirty-five in the eighth grade, in all one hundred fifty-two white boys. This made in all three hundred eighteen White boys.

The experiment was administered by Ikeda in June, 1928, in the schools of Hirosaki, Japan. All the Japanese boys were full blood Japanese. The district in which they live is off the beaten path of civilization and is one of the backward sections of Japan. Most of the children tested could not write English and their Japanese inscriptions had to be translated by Ikeda. There are no Ainus in or around this part of Japan, resident or traveling. Ainus can be found only in certain quarters of Hokaido which is separated from the main island, Hondo, by Tsugaru Strait. It is possible that some of the children here tested may have some trace of Ainu blood, however, which was mixed thousands of years ago.

It may be added that Ikeda had a thorough training in administration of tests in America.

Since space must be economized we present one curve for each grade group. These curves were derived by determining per cent of total work accurately done by each student and are called the individual curves. Each curve shown in the figure is the average of these individual curves. The curves for the Whites which are not here given were similarly derived from boys in the Virginia public schools.

An examination of the table will show that the Whites are some older than the Japanese and that they actually did more work in the time allotted than the Japanese

boys, the older Whites doing on the average 17.6 more columns and the younger ones 1.1 more.

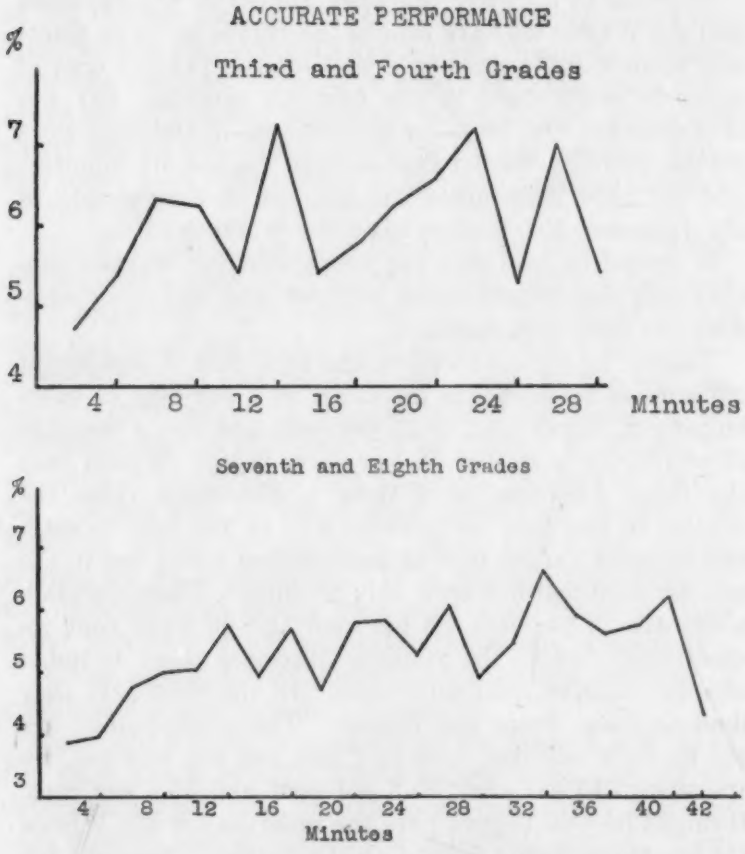
In order to compare the work curves of the Japanese and the Whites we have tapped the curves in three places and made distributions on a scale of the (1) per cent of accurate work done in the first six minutes, (2) per cent done the first half—or thereabout—of the total work period, and (3) the per cent done in the last six minutes; and we have determined the amount of overlapping of the Japanese distribution upon the White median.

It should be said that the young children worked continuously for twenty-eight minutes and the older children for forty-two minutes.

Taking into consideration the fact that a significant measure of difference is indicated by a per cent of overlapping of 25 per cent, or 75 per cent, and that a measure of similarity is one of 50 per cent, it will be seen that the older Japanese boys show a difference from the Whites in the first six minutes and in the first twenty-two minutes (about half of the working time) but in the last six minutes they tend only to differ. These overlappings are 23 per cent, 23 per cent and 60.6 per cent respectively. As to the younger Japanese there is not a case of positive, real difference. In the first half they tend to differ from the Whites. The overlappings are for first six minutes, first half and last six minutes, respectively, 45 per cent, 33.3 per cent and 56.8 per cent. It might be well to give here the medians for the Whites for the three points tapped on the curve. These are for younger Whites: 20.8 per cent, 51.3 per cent and 20.9 per cent; for older Whites: 15.5 per cent, 54.6 per cent (for first twenty-two minutes) and 13.8 per cent.

FIGURE 1

*Japanese Work Curves*




There are various observations that can be made upon examining the curves themselves but such are hardly profitable in the light of what is known about the overlapping of the racial distributions.

It is a fact that the older Japanese show more warming up than the older Whites but this contrast is not found in the case of the younger Japanese when they are compared with the younger whites. The older Whites

TABLE I

*Showing Various Comparisons of Japanese Boys with White Boys*

| No. | Age Yrs. | | Per Cent 1st 6 min. | | Per Cent 1st 14 min. | | Per Cent last 6 min. | | Tot. Acc. Columns | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Med. | Q. | Med. | Q. | Med. | Q. | Med. | Q. | Med. | Q. |
| 102 Young Japanese | 9.6 | .6 | 19.9 | 5.6 | 48.9 | 6.1 | 21.6 | 6.1 | 22.0 | 3.6 |
| 104 Older Japanese | 13.8 | .7 | 12.8 | 3.2 | 51.0 | 3.9 | 14.6 | 3.1 | 63.3 | 19.5 |
| Overlap on White Median* | | | | | | | | | | |
| Young Japanese | 9.0% | | 45.1% | | 33.3% | | 56.8% | | 55.0% | |
| Older Japanese | 18.8% | | 23.0% | | 23.0% | | 60.6% | | 24.0% | |

* Younger White, No. 161, Med. Age 10.6 yrs., Q. .9 yrs., Tot. Col. Med. 23.1, Q. 15.8.
Older White, No. 153, Med. Age 14.8 yrs., Q. 1.1 yrs., Tot. Col. Med. 80.9, Q. 22.8.

seem to get down to work quicker than the older Japanese but this is not the case with the younger Japanese. The older Whites seem to fatigue more than the older Japanese but the case is not so clear for the younger Japanese and younger Whites.

In the light of such findings we are disposed to think that this is not a case of racial differences in work curves but that such factors as selection and nurture have entered to make the differences as found in the older groups.

## BIBLIOGRAPHY

1. Garth, T. R., "Racial Differences in Mental Fatigue." *Journal Applied Psychology*, 1919, pp. 235-244.

2. Garth, T. R., "White, Indian and Negro Work Curves," *Journal Applied Psychology*, 1920, pp. 14-25.

3. Garth, T. R., "Mental Fatigue of Mixed- and Full-Blood Indians," *Journal Applied Psychology*, 1922, pp. 331-341.

4. Garth, T. R., "Mental Fatigue of Indians of Nomadic and Sedentary Tribes," 1926, *Journal Applied Psychology*, 437-452.

5. Garth, T. R., Walter Holcomb and Irma Gesche, "Mental Fatigue of Mexican Children," *Journal of Applied Psychology*, 1932, pp. 675-680.

# MULTIPLYING THE POSSIBILITIES OF THE MULTIPLE CHOICE FORM OF OBJECTIVE QUESTION

NORMA V. SCHEIDEMANN
*University of Southern California*

To the writer the most satisfactory type of objective question, especially for college people when it is desirable to approach a question from many angles, is the multiple choice. A common method of procedure with this type of question is to offer four choices for the completion of a statement, three wrong completions and one right completion, and to ask the student to select the correct one. This method was followed by the writer in a recent examination when a realization that the value of this type of question could be increased four-fold was thrust upon her. Examination forms were given to students with the directions that they should indicate the correct completion for each statement. As the examination progressed the students' repeated questions "Is only one answer correct in question——?" indicated to the writer that had the directions been to "specify all the completions to the given statements that you think are correct" the value of the examination would have been increased four-fold.

A four-fold increase in the value of a four-response multiple choice question is possible only when the questions are constructed properly. Questions that offer four choices for their completion do not, necessarily, represent a four-fold test of that question. The following two questions illustrate this:

**A neurological explanation of special reading disability points out that the two hemispheres of the brain have (1) different structure, (2) similar functions, (3) different size and design, (4) reversed patterns.**

**A neurological explanation of special reading disability points out that the two hemispheres of the brain have (1) similar structure, (2) different structure, (3) similar functions, (4) different functions.**

The first form tests four phases of the question, each challenging the student to a "yes-no" evaluation. The second form tests only two phases of the question, each of which is restated in what is equivalent to a negative statement of one of the phases. A "yes-no" evaluation of the two phases automatically answers thc negative statements of those phases.

When four-choice response questions are prepared properly and when the student is instructed to select only one answer of the four given, the student may be tested on only one phase of the given question. That is, he may be certain one completion is correct and uncertain in regard to the other three. If he is assured that only one completion is correct he will receive the same credit as if he were certain that the other three completions are faulty. If a student, however, is asked to indicate all of the correct answers and has no idea how many may be correct he will give objective evidence in regard to his knowledge of each completion (barring accidental success).

The following two questions illustrate the point just considered:

**A study of the nature of errors children make in reading tests will not determine (1) whether alternate sinistrad and dextrad eye-movements impede reading, (2) the severity of the disability, (3) or give clues to special methods in retraining, (4) specific remedial methods that are satisfactory to the exclusion of all other methods.**

**In comparison with day schools, the deaf pupilage of residential schools has (1) a predominant percentage of girls, (2) a higher percentage of foreign-born, (3) a higher percentage of congenitally deaf, (4) a higher percentage whose onset of deafness occurred at a later age.**

Examination of these questions will disclose that each presents opportunity to evaluate four phases of a topic.

These phases are non-overlapping. In the last illustration the student may have a knowledge of the relative percentage of girls and boys in our deaf pupilage, but may have no knowledge of the percentage of foreign-born pupilage, or congenitally deaf pupilage, or age at onset. In the preceding question the student may know the relation of eye-movements and reading tests but may not know about the factors included in the other three phases of the question. Informing the student that only one response is correct and that he is to select only one, fails to test him on three other phases. By not informing the student in regard to the correct number of responses offered and asking that he select all the correct ones, the student is given four times as great an opportunity to indicate his knowledge in regard to this topic.

In order to obtain objective data in regard to this subject the writer prepared forty-four choice response questions as a part of a final examination for an advanced course in the Psychology of Exceptional Children. The students were given the questions with the following instructions: "You have a number of four-choice response questions. For each question, indicate all the responses that are correct. Only one response may be correct, or two, three, or even all the responses may be correct. Write down the numerals of *all* the correct responses." When the students had completed this part of the examination they were asked to hand in their papers. It was then explained to the students that while the instructions were that any number of responses might be correct, only one response for each question actually was correct. The class was then asked to indicate the one correct response for each question. The students did not have access to their first answers during this part of the examination.

In the present experiment sixteen students understood and followed directions. A few students recognizing the customary form of the questions did not listen to

the directions; their papers had to be disregarded when the results of the experiment were tabulated.

The range of wrongs when the students were uninformed as to how many responses were right was from 5 to 40. When the students were informed that there was only one correct response for each question, the range was from 3 to 19. In each case the number of errors per student showed a tendency toward a normal distribution, but only three students, the best and the two poorest, retained their relative positions in the distributions. A grouping of the students into quartiles for each test showed that ten students retained positions within the same quartile for both tests. When the information that only one response for each question was correct was given, three students fell to a lower and three went to a higher quartile. Of the cases who remained within the same quartiles two occurred in the first, one in the second, three in the third, and four in the fourth.

The number of cases in this experiment were too few to warrant generalizations but for this class it seems that there were advantages in not informing students how many of several responses to a question were correct. The advantages indicated by this small class seem to be the same as those obtained from lengthening a test. Finer gradations resulted so that the recognition of distinctions in achievement was facilitated for the instructor.

A more desirable and more practical form of the multiple response question is suggested by this experiment. In constructing the questions some may be constructed with more than one correct response. This does not complicate scoring unduly and it makes possible the repeated use of this form of question for obviously, except for experimental purposes, the student should not be informed that more than one response may be correct when in every case only one response actually is correct.

## NEWS AND NOTES

Former United States Commissioner of Education, Dr. William John Cooper, has recently directed a national survey of secondary education which is to be published in a series of twenty-eight monographs. A special phase of the survey deals with changes in the social and economic order which reveal a more significant need of bringing counsel and guidance to boys and girls. Four outstanding guidance programs may be mentioned: (1) Centralized bureaus of guidance for secondary schools in city systems, as in Boston, Chicago and Cincinnati; (2) city school systems with a central guidance organization but with the individual secondary school considered the unit in the program; (3) centralized bureaus or departments in individual secondary schools; (4) and central guidance organizations in individual secondary schools which utilize regular officers and teachers as guidance functionaries. It is also disclosed that comparatively few schools are making thorough provisions for individual differences.

---

A meeting of the representatives of the Psychological Corporation will be held Friday, September 8, 1933, 2 P. M., at the Social Science Building, University of Chicago. This meeting, open to all Psychologists and invited business men, will be in charge of President Walter Dill Scott, Northwestern University. The speakers will include Mr. Lee H. Bristol, Vice President of the Bristol Myers Company, and Dr. S. N. Stevens, Northwestern University. The theme of the lectures will be Scientific Methods in Consumer and Market Research, with discussion led by Dr. Henry C. Link, Director of Market Research for the Psychological Corporation.

---

The American Psychological Association meetings will be held from Thursday to Wednesday, September 7-13, 1933, at the University of Chicago. The length of the period during which the meetings will be held has been extended to a full week with most of the sessions scheduled for mornings only in order to allow members ample opportunity to attend the Century of Progress Exposition. Further information may be obtained from the Secretary, Dr. Donald G. Paterson, University of Minnesota, Minneapolis, Minn.

---

The following announcements have come from Stanford University: Dr. E. R. Hilgard, instructor at Yale University, is to be assistant professor of Psychology at Stanford University. Dr.

Quinn McNemar has received a Social Science fellowship for further work in mathematical statistics at Chicago, Columbia and Harvard; Mr. Malcolm A. Campbell has accepted an instructorship at New York University; Mr. Michael Tomilin has been appointed research assistant in the Yale Laboratory of Comparative Psychology with assignment to the Anthropoid Experiment Station at Orange Park, Florida.

---

A study of foster Indian children in white homes is being sponsored by the University of Denver, under the guidance of Dr. Thomas R. Garth, professor of experimental psychology. Each child who is being studied is given a series of standard intelligence tests to determine whether his white environment is able to bring his intelligence as measured up to that of a normal white child. This is a phase of the nature-nurture problem. This experiment will also help answer the question: Are there really differences in racial traits? For if environment can wipe out the differences, then the differences are not really native.

---

Word has been received from The National Home Library Foundation in Washington that Professor Albert Einstein has consented to serve on its Advisory Board. The famous scientist together with the American scientist, Robert A. Millikan, will suggest new titles for inclusion in the Jacket Library to develop a broader interest in science. The Jacket Library books are to be distributed to the entire membership of eight thousand local Granges.

---

The International Bureau of Education has recently issued a review of the seven general principles of the teaching program of studies for infants and primary schools adopted by a decree of the Council of State of Switzerland. The first of these principles aims at giving "indispensable instruction and moral education to all," with a view to the scholars' future career, their intellectual, moral and social life, and to help them attain the maximum development they are capable of acquiring, under the most favorable conditions.

---

The Carnegie Foundation for the Advancement of Teaching has issued a summary of the current Review of Legal Education in the United States and Canada for the year 1932. Among other interesting items it states the effect of the depression upon law school attendance, which shows a loss of 6,980 students, or 15 per cent, between 1928 and 1931. Current figures are about 2½ per cent below those of 1931. The Review also records changes made dur-

ing the past year in bar admission requirements and outstanding developments among the law schools. Comparative tables show the present requirements for admission to the bars of each of the sixty states and Canadian provinces, and changes in the number of law schools of different types, and of their students, during the past forty years. Copies of this Review and other publications may be had without cost by addressing the Foundation, 522 Fifth Avenue, New York City.

---

Thousands of teachers and school executives will find it impossible this year to attend summer school due to the depression. But Research University, Washington, D. C., is meeting this need of continued growth by offering a new type of home study course, including over a hundred subjects, to graduates and undergraduate, teachers and others, at half the pre-depression rate of tuition. Also to those who can afford to spend six weeks in Washington the University offers opportunity to study in residence at the low cost of five dollars a week for tuition, with other expenses in proportion. Those interested may obtain further information by writing to Research University, 1607 Eye Street, Washington, D. C.

---

Nineteen thirty-three is the hundredth anniversary of higher education for women in the United States. In a review of a century of progress in college education for women the Federal Office of Education makes the above announcement. Oberlin College is the oldest coeducational college in the United States, although during its first four years it was coeducational only in the preparatory department. "Four women had sufficiently prepared to enter college work by 1837 and were graduated in 1842. By 1890 one out of every three college students was a woman, and the total enrollment of all college students was 156,756. By 1930 there were seven times as many college students—about 1,085,000. Nearly one-half of all students in the colleges now are women."

## BOOK REVIEWS

TAUSSIG, F. W. and JOSLYN, C. S., *American Business Leaders*, New York, MacMillan, 1932, pp. xiv + 319.

The writers set out to determine the social classes from which American business leaders are recruited; the contribution of each class compared with its proportion in the population as a whole and finally the relative importance of heredity and environment in the constitution of business leaders. The inquiry is apparently part of a larger philosophical problem as to the comparative importance of ability and opportunity. The authors selected business rather than the professions because the man without a college education is not particularly handicapped in the former.

The method consists essentially of a questionnaire to a group of business leaders selected from Poor's Register of Directors. Something over 7,000 returned questionnaires constitute the main sample. Psychologists will be interested in the procedure used to secure adequate returns. The covering letter headed from the Department of Economics at Harvard University had likewise in the heading an advisory committee who had been induced to lend their moral support to the undertaking. Names such as Julius H. Barnes and Owen D. Young appeared in the advisory list. As a result about half of the questionnaires were returned. Information was involved as to present position, age at which it was assumed, previous training, father's and grandfather's occupation, whether relatives or friends were interested in the concern and whether one received outside financial aid during the early stages of his business career. The methods used in deciding what blanks were too inadequately executed to use and in making numerous analyses of the data are outlined in considerable detail. Evidently the job was very carefully done from a technical standpoint.

Returning to the first main problem as to the special class from which the business leaders were recruited we find that 12 per cent had fathers who were farmers, 10 per cent laborers, and 57 per cent business men. In comparison with these figures we note that at about the era when the parents in question would be active, the laboring class constituted some 45 per cent of the employed population and the business group only about 10 per cent. Consequently the contribution of business parents to the present group of business leaders was still more striking.

The question then arises as to whether the foregoing differences are traceable to privilege or to superiority. The authors lean to the latter alternative. They find, for example, that the pro-

portion of laborers' sons among business leaders is no greater than that in other occupations, such as the professions where there are more rigid requirements, such as a college education. The results indicate further that the majority of the individuals did not receive help from relatives or friends nor were the more successful more favored in this way. It did not appear that the business leaders from the lower classes gave evidence of particular ability, that is, had to surmount a handicap by greater innate ability. Neither did it seem that those who came from the lower classes took a significantly longer time to attain success. The weak point in the argument which the authors see themselves is the possible rôle of other factors such as the social atmosphere surrounding the youngster in the business home. The authors feel, however, that if these factors were very pronounced the sons of laborers if they came through at all should have achieved greater success on the whole, which they do not do. They indicate further that as far as suggestion goes, there is so much publicity involved in newspapers and in general reading to the effect that people should try to make money, that the additional incentive given in the business home would not add very much to the child's attitude anyway. Furthermore, the average American boy even of poor parents can get enough schooling to enter successfully upon a business career. The authors feel consequently that these multiple factors do not loom very large in the picture. The reviewer's feeling would be that perhaps the attitude and atmosphere in the home may loom a little larger than the authors feel and that their conclusions should be tempered somewhat but he is not disposed to question their main thesis except by way of toning it down slightly. Their final conclusions on the matter may be best expressed by quoting the last sentence of the book: "Our results strongly suggest, even if they do not prove, that inequality of earning between the several occupational classes has its origin in a fundamental inequality of native endowments rather than in an inequality of opportunities."

The book is well written and systematic and evidently was a careful piece of work. It represents a rather novel approach to these broad questions of heredity versus environment and psychologists interested in those problems will find it of considerable value.

SCHMALHAUSEN, SAMUEL D., Editor. *Our Neurotic Age. A Consultation.* Murray Hill, New York. Farrar and Rinehart, Inc.

The title indicates that the author conceived neuroticism on a grand scale. For clarity the articles by twenty-seven authors are arranged into four divisions according to their slant. In Part I,

Is the Normal Mind Sane? appear the contributions of Morgan, Roback, Jones, Jastrow, et al. Morgan, in The Mind in the Breaking, throws into relief the struggles incident in life and the equipment with which we must meet our difficulties. Experience results in attitudes or emotional habit patterns. These must be integrated if we are to fight a winning battle. If they are not integrated we are at odds with ourselves and will likely break before we can overcome the external difficulty. Mental balance depends upon the attitude which we assume toward the thing we do not understand. There are two principles which should be applied in all child training: teach the child to face the issues of life frankly, and teach him to react to them objectively. Roback, "the most erudite prince," in "What is Sanity?" points out that sanity includes normality, rationality (that rare ability of representing things as they are), a coordination between the volitional and the critical faculties, emotional stability, and an impulse to carry out one's convictions. He predicts a test which will measure the relative sanity of "sane" individuals.

Ernest Jones, in The Concept of a Normal Mind, touches several ideas in rather a sketchy manner. There is need of a standard of normality because of variations. A statistical normality, however, cannot be considered as synonymous with a psychologically healthy mind for the community cannot be accepted as capable of developing the best there is in the mind. There are three attributes of normality: happiness, efficient mental functioning, and a positive social feeling of relationship. The last (friendliness) is measured, not objectively but by an internal freedom from repressions. The second is obtained by an unimpeded flow of energy from the primordial unconscious sources to response. The first one is defined as the capacity for enjoyment plus self content. Impairment of happiness is due to the triad of fear, hate and guilt, of which the most important is fear. Libidinal tensions lead to defenses of hate and guilt. The psychological problem of normality must reside in the ability to hold wishes in suspension, without renouncing, or reacting to, them in defensive ways. The nearest attainable criterion of normality is fearlessness.

Jastrow, in "The Maladies of Thought," points out two orders of thought: "The autonomic component represents the primary function. Within its sovereignty the glandular stresses and their affective radiations exert their sway; upon this is imposed the cerebro-spinal direction. Within the latter sphere of influence appears gradually, sharply accentuated in man, the cortical dominance . . .

In Part II, Beyond Normality, one feels that his mind has slipped into a manhole for a sixteen mile jaunt in Freudian sewage.

One doubts whether the idea ever occurred to some of the authors that character is more essential in our society today than the avoidance of every repression.

Parts III and IV in a general way express some of the ideas of the editor in the Introduction: "The social system is admitted to be in the last stages of a wasting disease . . . Is modern man neurotic or rather modern civilization? . . . The evidence at hand is overwhelmingly in favor of the thesis that our competitive civilization is the fertile begetter of psychoneurotic malaise and misery . . . In the Soviet Republic . . . mental hygiene is synonymous with the utter socialization and humanization of everyday life. Communism is a great contribution to psychoanalytic therapy, sociologically envisaged, in the realistic sense that it is able to transform the masses of 'nonentities' . . . into human entities . . ." This sounds like a day dream.

In this volume one meets the King and the cat. Part I is quite worth while and needs no further comment. The editor seems to have quite a journalistic flair but this effort in psychology is somewhat bizarre. Parts of this volume suggest rather crude forms of vicarious exhibitionism; much of it could be regarded as a tramp through psychological slums. There seems to be an attempt to make Hollywood seem the normal state of health of American society rather than the abscess that it is. Is this a compensatory volume or is it meant to serve as a safety valve for some neurotic conditions?

JOHN R. GENTRY,
*Ohio University.*

---

## NEW BOOKS AND PAMPHLETS RECEIVED

Books and pamphlets for review should be sent to James P. Porter, Editor, Journal of Applied Psychology, Ohio University, Athens, Ohio.

*Articulation of High School and College.* P. ROY BRAMMELL. United States Department of the Interior, Bulletin No. 17, 1933, Monograph No. 10. Washington, D. C. 10 cents. 96 pp.

*The Blind in School and Society.* THOMAS D. CUTSFORTH. D. Appleton and Company, New York City, 1933. $2.50, 263 pp.

*Do College Students Choose Vocations Wisely?* EDWARD J. SPARLING. Bureau of Publications, Teachers College, Columbia University, New York City, 1933. $1.50. 110 pp.

*The Dynamics of Therapy.* JESSIE TAFT. Macmillan Company, New York City, 1933. $2.50. 296 pp.

*Health Work and Physical Education.* Monograph No. 28, Bulletin No. 17, 1932; *Intramural and Interscholastic Athletics.* Monograph No. 27; P. Roy Brammel. National Survey of Secondary Education, U. S. Government Printing Office, Washington, D. C.

*A Point Scale of Performance Tests:* Vol. I. Clinical Manual. $1.50; Vol. II, *The Process of Standardization.* Grace Arthur. Commonwealth Fund, New York City, 1933. $1.50. 106 pp.

*Psychological Tests, Methods and Results.* Henry E. Garrett and Mr. R. Schneck. Harper & Brothers, New York City, 1933. 235 pp.

*Psychology.* Frederick H. Lund. Ronald Press Company, New York City, 1933. $3.00. 475 pp.

*Psychology for Teachers.* Revised Edition. Charles E. Benson, James E. Lough, Charles E. Skinner and Paul V. West. Ginn & Company, New York City, 1933. 490 pp.

*Revista de Psicologia i Pedagogia,* Vol I, No. 1, February, 1933. L'Institut Psicotecnic de la Generalitat i el Seminari de Pedagogia de la Universitat, Barcelona, Spain.

*Scholastic Aptitude Tests for Medical Students.* F. A. Moss. Journal of Association of American Medical College, January, 1933. 27 pp.

*Die Seelische Eingliederung in Die Gesellschaft.* Wilhelm Steinberg. Verlag Ernst Reinhardt, Muchen. 127 pp.

*Seven Psychologies.* Edna Heidbreder. The Century Company, New York City, 1933. $2.25. 450 pp.

*Specialization of Verbal Facility at the College Entrance Level.* Warren George Findley. Bureau of Publications, Teachers College, New York City, $1.50. 76 pp.

*University of Kansas Studies in Psychology, No. 1.* R. B. Wheeler, Editor. Psychological Monograph, Psychological Review Company, Princeton, N. J., and Albany, N. Y. 300 pp.

*Wirkungsweisen von Farbeindrucken auf Schwachsinnige Kinder.* Max Wittman. Forschungen uber Farbspielkunst, Vol. I, No. 2. Verlag Von S. Hirzel, Leipzig, Germany. 122 pp.

## Directory of American Psychological Periodicals

*(Continued)*

**Psychological Abstracts**—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $6.00. 700 pages ann. Edited by W. S. Hunter. Monthly. Abstracts of psychological literature. Founded 1927.

**Pedagogical Seminary and Journal of Genetic Psychology**—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Child behavior, animal behavior, and comparative psychology. Founded 1891.

**Journal of Social Psychology**—Worcester, Mass.: Clark University Press. Subscription $7.00. 500-600 pages annually. Edited by John Dewey and Carl Murchison. Quarterly. Political, racial, and differential psychology. Founded 1929.

**Personnel Journal**—Baltimore: The Williams & Wilkins Co. Subscription $5.00. 500 pages, approximately. Edited by Walter V. Bingham. Bi-Monthly.

**Mental Measurement Monographs**—Baltimore: The Williams & Wilkins Company. Subscription $5. 400 pages annually. Edited by Knight Dunlap and Buford Johnson. Published without fixed dates, each number a single research. Founded 1925.

**Studies in Psychology and Psychiatry**—Baltimore: The Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Edward A. Pace. Published without fixed dates, each number a single research. Founded 1926.

**Journal of Philosophy**—New York: 515 West 116th Street. Subscription $4.00. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush, and H. W. Schneider.

**Child Development**—Baltimore, Md.: Williams & Wilkins Company. Subscription $5.00. 400 pages annually. Edited by Buford Johnson. Quarterly. Founded 1930.

**Character and Personality**—Durham, N. C.: Duke University Press. Subscription $2.00 per annum. Quarterly. Edited by Robert Saudek (London) and Ernest Seeman (Durham, N. C.). Founded 1932.

# Directory of American Psychological Periodicals

**American Journal of Psychology**—Ithaca, N. Y.: Cornell University. Subscription $6.50. 624 pages annually. Edited by M. F. Washburn, K. M. Dallenbach, M. Bentley, E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

**Journal of General Psychology**—Worcester, Mass.: Clark University Press. Subscription $14.00 per annum; $7.00 per volume. 1,000 pages annually (2 volumes). Edited by Carl Murchison and an international cooperating board. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

**Psychological Review**—Princeton, N. J.: Psychological Review Company. Subscription $5.50. 540 pages annually. Bi-monthly. General Psychology. Edited by Howard C. Warren. Founded 1894.

**Psychological Monograph**—Princeton, N. J.: Psychological Review Company. Subscription $6.00 per vol. 500 pp. Ed. by Raymond Dodge. Published without fixed dates, each issue one or more researches. Founded 1895.

**Psychological Index**—Princeton, N. J.: Psychological Review Company. Subscription $4.00. 300-400 pp. Edited by W. S. Hunter. An annual bibliography of psychological literature. Founded 1895.

**Psychological Bulletin**—Princeton, N. J.: Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Edited by Edward S. Robinson. Founded 1904.

**Archives of Psychology**—New York: Columbia University. Subscription $6. 500 pp. ann. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study. Founded 1906.

**Journal of Abnormal and Social Psychology**—Eno Hall, Princeton, N. J. American Psychological Association. Subscription $5. Foreign, $5.25. Editor, Henry T. Moore. Quarterly. 448 pages annually. Abnormal and Social. Founded 1906.

**Psychological Clinic**—Philadelphia: Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene. Founded 1907.

**Psychoanalytic Review**—Washington, D. C.: 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Ed. by W. A. White and S. E. Jelliffe. Founded 1913.

**Journal of Experimental Psychology**—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Bi-monthly. Edited by S. W. Fernberger. Founded 1916.

**Journal of Applied Psychology**—Indianapolis: The C. E. Pauley & Co., Inc. Subscription $5.00. 400 pages annually. Bi-monthly. Edited by James P. Porter, Ohio University, Athens, Ohio. Founded 1917.

**Journal of Comparative Psychology**—Baltimore: The Williams & Wilkins Co. Subscription $5.00 per volume of 450 pages. Three volumes every two years. Edited by Knight Dunlap and Robert M. Yerkes. Founded 1921.

**Comparative Psychology Monographs**—Baltimore: Johns Hopkins University Press. Subscription $5. 400 pages per volume. Knight Dunlap, Managing Editor. Published without fixed dates, each number a single research.

**Genetic Psychology Monographs**—Worcester, Mass.: Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Edited by Carl Murchison. Monthly. Each number one complete research. Child behavior, animal behavior and comparative psychology. Founded 1925.

*(Continued on Third Cover)*